GENERALIZED LIMITS IN GENERAL ANALYSIS, FIRST PAPER*

BY

CHARLES N. MOORE

The analogies that exist between infinite series and infinite integrals are well known and have frequently served to indicate the extension of a theorem or a method from one of these domains of investigation to the other. According to a principle of generalization that has been formulated by E. H. Moore, the presence of such analogies implies the existence of a general theory which includes the central features of both the special theories.† It is the purpose of the present paper to develop the fundamental principles of that section of this general theory which contains as particular instances the theories of Cesàro and Hölder summability of divergent series and divergent integrals. Furthermore, the usefulness of the theory will be illustrated by proving a general theorem in it which includes as special cases the Knopp-Schnee-Ford theorem‡ with regard to the equivalence of the Cesàro and Hölder means for summing divergent series, an analogous theorem due to Landau § concerning divergent integrals, and a further new theorem with regard to the equivalence of certain generalized derivatives.

The general theorem just mentioned can be extended to the case of multiple limits so as to include other new theorems, analogous to those referred to above, with regard to multiple series, multiple integrals, and partial derivatives. This extension, however, involves formulas that are considerably more complicated than in the case of simple limits. I shall therefore reserve it for a second paper, as I wish to avoid algebraic complexity in this first presentation of the general theory.

Following the terminology introduced by E. H. Moore,|| we indicate the basis of our general theory as follows:

$$(\mathfrak{A};\ \mathfrak{P};\ \mathfrak{S};\ \mathfrak{G}^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}};\ \mathfrak{H}^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}};\ \mathfrak{F}^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}};\ \phi_0;\ J^{\text{on } \mathfrak{G} \text{ to } \mathfrak{H};\ \text{on } \mathfrak{H} \text{ to } \mathfrak{F}})$$

* Presented to the Society, December 28, 1918.

† Cf. E. H. Moore, *Introduction to a Form of General Analysis*, The New Haven Mathematical Colloquium, Yale University Press, 1910, p. 1.

‡ That the existence of the Hölder limit implies the existence and equality of the Cesàro limit of the same order was first proved by Knopp; cf. his Inauguraldissertation, *Grenzwerte von Reihen bei der Annäherung an die Konvergenzgrenze*, Berlin, 1907. The converse theorem was established independently by Schnee and W. B. Ford; cf. Mathematische Annalen, vol. 67 (1909), pp. 110–125, and American Journal of Mathematics, vol. 32 (1910), pp. 315–326.

§ See Leipziger Berichte, vol. 65 (1913), pp. 131–138.

|| Cf. his two papers: *On the foundations of the theory of linear integral equations*, Bulletin

where $\mathfrak{A} \equiv [a]$ denotes the class of all real numbers a, $\mathfrak{P} \equiv [p]$ denotes a class of elements p, and $\mathfrak{S} \equiv [\sigma]$ denotes a class of sets σ of elements p of the range $\mathfrak{P}$; $\mathfrak{G} \equiv [\gamma]$, $\mathfrak{H} \equiv [\eta]$, and $\mathfrak{F} \equiv [\phi]$ are three classes of functions γ, η, and ϕ respectively on $\mathfrak{S}$ to $\mathfrak{A}$ (we shall restrict ourselves throughout to the consideration of single-valued functions); ϕ_0 is a special function ϕ of the class $\mathfrak{F}$; and J is a function on $\mathfrak{G}$ to $\mathfrak{H}$ and on $\mathfrak{H}$ to $\mathfrak{F}$, that is a functional transformation turning a function of the class $\mathfrak{G}$ into a function of the class $\mathfrak{H}$ or a function of the class $\mathfrak{H}$ into a function of the class $\mathfrak{F}$, denoted by $J\gamma$ or $J\eta$.

In order to make clear the relationship of our general theorem to the two special theorems referred to above we will indicate here what the general basis reduces to in the particular instances III and IV:

$$\mathfrak{P}^{\mathrm{III}} \equiv [\text{all } n = 1, 2, 3, \cdots]; \qquad \mathfrak{S} \equiv [\sigma_n \equiv (1, 2, \cdots, n) | n];$$

$$\mathfrak{G} \equiv \mathfrak{H} \equiv \mathfrak{F} \equiv [\text{all } \gamma, \eta, \phi^{\text{on } \mathfrak{S} \text{ to } \mathfrak{A}}]; \qquad \phi_0(\sigma_n) = n \ \ (n);$$

$$(J\gamma)(\sigma_n) = \gamma(\sigma_1) + \gamma(\sigma_2) + \cdots + \gamma(\sigma_n) \ \ (n);$$

$$(J\eta)(\sigma_n) = \eta(\sigma_1) + \eta(\sigma_2) + \cdots + \eta(\sigma_n) \ \ (n);$$

$$\mathfrak{P}^{\mathrm{IV}} \equiv [\text{all } a > 0]; \qquad \mathfrak{S} \equiv [\sigma \equiv (\text{all } x \text{ such that } 0 < x \leqq a) \ \ (a > 0)];$$

$\mathfrak{G} \equiv$ [all functions that are finite and integrable (Lebesgue) on every finite interval $(0 < x \leqq a)$];

$$\mathfrak{H} \equiv \left[\text{all } \eta = \left(\int_0^x \gamma | x > 0 \right) \right]; \qquad \mathfrak{F} \equiv \left[\text{all } \phi = \left(\int_0^x \eta | x > 0 \right) \right];$$

$$\phi_0(\sigma_a) = a \ \ (a); \quad (J\gamma)(\sigma_a) = \int_0^a \gamma(\gamma a); \quad (J\eta)(\sigma_a) = \int_0^x \eta(\eta a).$$

We next proceed to make certain postulates with regard to the nature of the elements in our basis, readily seen to be verified in the specific instances indicated. Thus we require the class $\mathfrak{S}$ to have the following properties:

(U) Either corresponding to every σ' there exists a least common superclass of classes $\sigma < \sigma'$, or there exists a σ_0 such that for every $\sigma' > \sigma_0$ there exists a least common superclass of classes $\sigma < \sigma'$. In both cases the least common superclass is itself a σ.

(A) Corresponding to every σ' there exists a least common subclass of classes $\sigma > \sigma'$, this least common subclass being itself a σ.

In the typical instances in view in the formation of this general theory, of the two alternatives in (U) one holds and the other does not hold; however, it is not assumed that this disjunction between the two alternatives shall be

of the American Mathematical Society, vol. 18 (1912), pp. 334–362; *On the fundamental functional operation of a general theory of linear integral equations*, Proceedings of the Fifth International Congress of Mathematicians, Cambridge, 1913, pp. 230–255.

presupposed. In order to avoid notational reference to these alternatives, it is convenient to introduce a property $-$ of sets σ of $\mathfrak{S}$; if the first alternative holds, every σ of $\mathfrak{S}$ has the property $-$, in notation $\bar{\sigma}$; if the first alternative does not hold, the sets $\bar{\sigma}$ are the sets $\sigma > \sigma_0$; further for brevity we introduce a property $\cdot$ (the negation of $-$); thus every σ is a $\bar{\sigma}$ or a $\dot{\sigma}$.

We now define

$$\sigma'_{-1} \equiv [\text{the least common superclass of classes } \sigma < \sigma'] \quad (\bar{\sigma}');$$

$$\sigma'_1 \equiv [\text{the least common subclass of classes } \sigma > \sigma'];$$

$$\sigma'_{n+1} \equiv [\text{the least common subclass of classes } \sigma > \sigma'_n] \; (\sigma | n = 1, 2, \cdots).$$

We then postulate

(R) Corresponding to every σ', σ'_1 is a $\bar{\sigma}$ and there exists $(\sigma'_1)_{-1} = \sigma'$.

We next define the notation $\sigma' < \sigma''$, $\sigma'' > \sigma'$, to mean that σ'' contains all the elements of σ' and at least one element not found in σ'. We are then ready to formulate three limit definitions which are based on the fundamental definition of limit in General Analysis given by E. H. Moore.*

For a given θ on $\mathfrak{S}$ to $\mathfrak{A}$, a given a, and a given σ' such that there exists $\sigma < \sigma'$, we shall write

$$\lim_{\sigma | \sigma < \sigma'} \theta(\sigma) = a$$

in the case that, corresponding to an arbitrary positive number e, there exists a $\sigma_e < \sigma'$ such that for every σ having the property $\sigma_e \leqq \sigma < \sigma'$, $|\theta(\sigma) - a| < e$.

For any function θ on $\mathfrak{S}$ to $\mathfrak{A}$ we shall say that $\theta(\sigma)$ approaches a limit as to σ if corresponding to every positive e we can find a σ_e such that for every $\sigma > \sigma_e$ we have $|\theta(\sigma) - a| < e$.

For any function θ on $\mathfrak{S}$ to $\mathfrak{A}$ we shall mean by the notation

$$\lim_{\sigma} \theta(\sigma) = \infty$$

that for every positive e there exists a σ_e such that for every $\sigma > \sigma_e$, $\theta(\sigma) > e$.

We define the notation

$$(D\theta)(\sigma) = \alpha(\sigma)$$

with regard to every θ, α on $\mathfrak{S}$ to $\mathfrak{A}$, to mean that

$$(1) \quad \lim_{\sigma | \sigma < \sigma'} \frac{\theta(\sigma') - \theta(\sigma)}{\phi_0(\sigma') - \phi_0(\sigma)} = \alpha(\sigma') \quad (\bar{\sigma}'), \qquad \theta(\sigma) = \alpha(\sigma) \quad (\dot{\sigma}).$$

*Cf. Proceedings of the National Academy of Sciences, vol. 1 (1915), pp. 628–632.

We require the class $\mathfrak{G}$ to have the linear property

(L) as defined by E. H. Moore,*

and the property (P) defined by

(P) The product $\gamma_1(\sigma)\cdot\gamma_2(\sigma)$ is a function of the class $\mathfrak{G}$.

It will then follow that the product $\gamma_1(\sigma)\cdot\gamma_2(\sigma)\cdot\ \cdots\ \cdot\gamma_n(\sigma)$ is a function of the class $\mathfrak{G}$.

We require the class $\mathfrak{H}$ to have the properties (L) and (P) and the further property of being a subclass of the class $\mathfrak{G}$, which property we designate as S_G. We further postulate for the class $\mathfrak{H}$ the property (B) defined by

(B) If $\lim\limits_{\sigma}\eta(\sigma)$ exists and is equal to a, then $|\eta(\sigma)| < a_1 \quad (\sigma)$.

We require the class $\mathfrak{F}$ to have the properties (L) and (P) and the further property of being a subclass of the class $\mathfrak{H}$, which property we designate as S_H. Hence J is also on $\mathfrak{F}$ to $\mathfrak{F}$. We also postulate for the class $\mathfrak{F}$ the property (Δ) defined by

(Δ) There exists $D\phi \equiv [(D\phi)(\sigma)|\sigma]$, a function of the class $\mathfrak{H}$.

We now define

$$(2) \quad \overline{\phi}(\sigma) \equiv \phi(\sigma_{-1}) \quad (\overline{\sigma}), \qquad \overline{\phi}(\sigma) \equiv 0 \quad (\dot{\sigma}); \qquad \dot{\phi}(\sigma) \equiv \phi(\sigma_1) \quad (\sigma);$$

and with regard to the functions $\overline{\phi}$ and $\dot{\phi}$ we postulate

(F) All functions $\overline{\phi}$ and $\dot{\phi}$ are of the class $\mathfrak{F}$.

We further postulate with regard to the class $\mathfrak{F}$

(C) For every ϕ and every $\overline{\sigma}'$ there exists $\lim\limits_{\sigma|\sigma<\sigma'}\phi(\sigma) = \overline{\phi}(\sigma')$.

For the operation J we postulate the following properties:

(M_1) If $a_1 < \gamma_1 < a_2$, $0 \leqq \gamma_2$, then

$$a_1(J\gamma_2)(\sigma) \leqq (J[\gamma_1\gamma_2])(\sigma) \leqq a_2(J\gamma_2)(\sigma) \quad (\sigma),$$

(M_2) If for every $\sigma > \sigma'$, $a_1 < \gamma_1 < a_2$, $0 \leqq \gamma_2$, then for $\sigma'' > \sigma'$

$$a_1[(J\gamma_2)(\sigma'') - (J\gamma_2)(\sigma')] \leqq (J[\gamma_1\gamma_2])(\sigma'') - (J[\gamma_1\gamma_2])(\sigma') \leqq a_2[(J\gamma_2)(\sigma'') - (J\gamma_2)(\sigma')],$$

(I_D) For every η and every σ there exists $(D(J\eta))(\sigma) = \eta(\sigma)$,

(I_J) For every ϕ and every σ there exists $(J(D\phi))(\sigma) = \phi(\sigma)$.

We next introduce for the sake of brevity the following notations:

$$(3) \quad \phi_{0n}(\sigma) = \phi_0(\sigma)\cdot\phi_0(\sigma_1)\cdot\phi_0(\sigma_2)\cdot\ \cdots\ \cdot\phi_0(\sigma_{n-1}) \quad (n > 1), \qquad \phi_{01}(\sigma) = \phi_0(\sigma).$$

* Loc. cit.

We then postulate with regard to ϕ_0

(I) ϕ_0 is a positive increasing function of σ,

(II) $\dfrac{1}{\phi_{0n}(\sigma)}(J^n\eta)(\sigma)$ as function of σ is of the class $\mathfrak{F}\quad(n)$,

the symbol $J^n\eta$ indicating that the operation J has been repeated n times,

(III) $\phi_0(\sigma_n)$ as function of σ is of the class $\mathfrak{F}\quad(n)$,

(IV) $\lim\limits_{\sigma}\phi_{0n}(\sigma) = \infty\quad(n)$,

(V) $[\phi_0(\sigma_1) - \phi_0(\sigma)]$ is constant for all σ,

(VI) $\phi_{0n}(\dot\sigma) = n\phi_{0,\,n-1}(\dot\sigma)\quad(n>1),\qquad \phi_{01}(\dot\sigma) = \phi_0(\dot\sigma) = 1.$

We have then as the foundation of our theory:

$$\Sigma \equiv (\mathfrak{A};\ \mathfrak{P};\ \mathfrak{S}^{UAR};\ \mathfrak{G}^{\text{on }\mathfrak{T}\text{ to }\mathfrak{A}\cdot LP};\ \mathfrak{H}^{\text{on }\mathfrak{S}\text{ to }\mathfrak{A}\cdot LPS_GB};$$
$$\mathfrak{F}^{\text{on }\mathfrak{S}\text{ to }\mathfrak{A}\cdot LPS_HC\Delta};\ \phi_0^{\mathfrak{A}\cdot\mathrm{I\ II\ III\ IV\ V\ VI}};\ \bar\phi^{\mathfrak{A}};\ \dot\phi^{\mathfrak{A}};$$
$$J^{\text{on }\mathfrak{G}\text{ to }\mathfrak{h}\cdot\text{on }\mathfrak{h}\text{ to }\mathfrak{z}\cdot M_1M_2I_DI_J}).$$

We will now prove that the operation J, when applied to the class $\mathfrak{H}$, has the linear property (L). Let us set

$$(J\eta_n)(\sigma) = \phi_n(\sigma)\qquad (n = 1, 2, \cdots, i).$$

It follows from the definition of D and I_D that

$$(D(a_1\phi_1 + a_2\phi_2 + \cdots + a_i\phi_i))(\sigma)$$
$$= a_1(D\phi_1)(\sigma) + a_2(D\phi_2)(\sigma) + \cdots + a_i(D\phi_i)(\sigma)$$
$$= a_1\eta_1 + a_2\eta_2 + \cdots + a_i\eta_i.$$

Applying J to both sides and making use of I_J, we have

$$(J(a_1\eta_1 + a_2\eta_2 + \cdots + a_i\eta_i)(\sigma)$$
$$= a_1(J\eta_1)(\sigma) + a_2(J\eta_2)(\sigma) + \cdots + a_i(J\eta_i)(\sigma).$$

We will now prove two properties of the operations D and J as applied to the class $\mathfrak{F}$. We have for $\sigma \neq \sigma'$

$$\frac{\phi_1(\sigma')\phi_2(\sigma') - \phi_1(\sigma)\phi_2(\sigma)}{\phi_0(\sigma') - \phi_0(\sigma)}$$
$$= \phi_2(\sigma')\frac{\phi_1(\sigma') - \phi_1(\sigma)}{\phi_0(\sigma') - \phi_0(\sigma)} + \phi_1(\sigma)\frac{\phi_2(\sigma') - \phi_2(\sigma)}{\phi_0(\sigma') - \phi_0(\sigma)}.$$

By virtue of properties (C) and (Δ) of class $\mathfrak{F}$ the right side of the above

equation approaches a limit as to $\sigma|\sigma < \sigma'$, for every $\bar{\sigma}'$. Hence so also does the left side, and we obtain the formula

$$(D[\phi_1 \phi_2])(\sigma) = \phi_2(\sigma)(D\phi_1)(\sigma) + \bar{\phi}_1(\sigma)(D\phi_2)(\sigma) \tag{4}$$

for $\sigma = \bar{\sigma}$. In view of (1) and (2) equation (4) obviously holds for $\sigma = \dot{\sigma}$.

By virtue of postulates (Δ) and (F) and the properties (P) and (L) of the class $\mathfrak{H}$ we may apply the operation J to equation (4). On doing so we obtain, in view of I_J and the linear property of J established above,

$$(J[\phi_2(D\phi_1)])(\sigma) = \phi_1(\sigma)\phi_2(\sigma) - (J[\bar{\phi}_1(D\phi_2)])(\sigma). \tag{5}$$

Equation (4) includes as special cases the formulas for differentiation of a product and forming the first difference of a product. Equation (5) includes as special cases the formulas for integration by parts and partial summation.

We shall next prove two further properties of the special function ϕ_0. The first of these properties is the following:

$$\text{(VII)} \quad (D\phi_{0n})(\sigma) = n\phi_{0,\,n-1}(\sigma) \quad (\sigma, n > 1), \qquad (D\phi_{01})(\sigma) = 1 \quad (\sigma).$$

We have, in view of the definition of D, (V), (R), (C), and (2),

$$\begin{aligned} (D\phi_{02})(\sigma') &= \lim_{\sigma|\sigma<\sigma'} \left[\phi_0(\sigma') \frac{\phi_0(\sigma_1') - \phi_0(\sigma_1)}{\phi_0(\sigma') - \phi_0(\sigma)} + \phi_0(\sigma_1) \frac{\phi_0(\sigma') - \phi_0(\sigma)}{\phi_0(\sigma') - \phi_0(\sigma)} \right] \\ &= \lim_{\sigma|\sigma<\sigma'} [\phi_0(\sigma') + \phi_0(\sigma_1)] = 2\phi_0(\sigma') \quad (\bar{\sigma}'). \end{aligned} \tag{6}$$

We now assume

$$(D\phi_{0i})(\sigma) = i\phi_{0,\,i-1}(\sigma) \qquad (i \geqq 2, \bar{\sigma}).$$

Then, making use of (4), we have

$$\begin{aligned} (D\phi_{0,\,i+1})(\sigma) &= (D[\phi_0 \dot{\phi}_{0i}])(\sigma) \\ &= \phi_0(\sigma)(D\phi_{0i})(\sigma_1) + \phi_{0i}(\sigma)(D\phi_0)(\sigma) \\ &= \phi_0(\sigma)[i\phi_{0,\,i-1}(\sigma_1)] + \phi_{0i}(\sigma) = (i+1)\phi_{0i}(\sigma) \quad (\bar{\sigma}). \end{aligned} \tag{7}$$

Hence, if (VII) holds for $\bar{\sigma}$, $n = i$ $(i \geqq 2)$, it will hold for $\bar{\sigma}$, $n = i + 1$. Combining this fact with equation (6), we infer that (VII) holds for $\bar{\sigma}$, $n \geqq 2$; for $\bar{\sigma}$, $n = 1$ it is an obvious consequence of the definition of D. For $\dot{\sigma}$, n it follows at once from the definition of D and (VI).

We now introduce the following notation:

$$\psi_n(\sigma) = \frac{1}{\phi_{0,\,n-1}(\sigma)} \qquad (n > 1). \tag{8}$$

The second property of ϕ_0 that we wish to prove may then be stated as follows:

$$\text{(VIII)} \qquad (D\psi_n)(\sigma) = -\frac{n-1}{\phi_{0n}(\sigma_{-1})} \qquad (\bar{\sigma}, n > 1).$$

We have from the definition of D, (R), (C), and (2)

$$(9)\quad \begin{aligned}(D\psi_2)(\sigma') &= \left(D\frac{1}{\phi_{01}}\right)(\sigma') = \lim_{\sigma\,|\,\sigma<\sigma'} \frac{\dfrac{1}{\phi_0(\sigma')} - \dfrac{1}{\phi_0(\sigma)}}{\phi_0(\sigma') - \phi_0(\sigma)} \\ &= \lim_{\sigma\,|\,\sigma<\sigma'} \frac{-1}{\phi_0(\sigma')\phi_0(\sigma)} = \frac{-1}{\phi_0(\sigma'_{-1})\phi_0(\sigma')} = -\frac{1}{\phi_{02}(\sigma'_{-1})} \quad (\bar{\sigma}).\end{aligned}$$

Then, assuming (VIII) for $n = i$, we have from (4) and (9)

$$(10)\quad \begin{aligned}(D\psi_{i+1})(\sigma) &= \left(D\left[\frac{1}{\bar{\phi}_0}\psi_i\right]\right)(\sigma_1) \\ &= -\frac{1}{\phi_{02}(\sigma_{-1})}\cdot\frac{1}{\phi_{0,\,i-1}(\sigma_1)} + \frac{1}{\phi_0(\sigma_{-1})}\cdot\frac{-(i-1)}{\phi_{0i}(\sigma)} \\ &= -\frac{i}{\phi_{0,\,i+1}(\sigma_{-1})} \quad (\bar{\sigma}).\end{aligned}$$

From (9) and (10) the proof by induction of formula (VIII) may readily be completed.

We are now ready to define the two generalized limits with which we shall be concerned. Given any function $\eta(\sigma)$, we set

$$(11)\quad (C_n\,\eta)(\sigma) \equiv [\,n!/\phi_{0n}(\sigma)\,](J^n\,\eta)(\sigma) \quad (n),$$

$$(12)\quad (M\eta)(\sigma) \equiv [\,1/\phi_0(\sigma)\,](J\eta)(\sigma),$$

$$(13)\quad (H_n\,\eta)(\sigma) \equiv (M^n\,\eta)(\sigma) \quad (n),$$

where $\phi_{0n}(\sigma)$ is defined as in equation (3) and C and H are used, as is customary, to connote Cesàro and Hölder. If for a fixed n $\lim_\sigma (C_n\,\eta)(\sigma)$ exists, we define this limit as the generalized limit of type (C_n) for $\eta(\sigma)$. If $\lim_\sigma (H_n\,\eta)(\sigma)$ exists, we define this limit as the generalized limit of type (Hn) for $\eta(\sigma)$.

We shall prove the equivalence of these two generalized limits. We begin by proving some lemmas.

LEMMA 1. *If we represent by E the identical functional operation $E\theta = \theta(\theta)$, we have the identity*

$$(14)\quad \left(\left(\frac{n-1}{n}M + \frac{1}{n}E\right)(C_n\,\eta)\right)(\sigma) = (M(C_{n-1}\,\eta))(\sigma) \quad (n),$$

where for the sake of uniformity we have set $(C_0\,\eta)(\sigma) = \eta(\sigma)$.

We have from the definition of $(C_n\,\eta)(\sigma)$

$$(15)\quad \begin{aligned}\phi_0(\sigma)\cdot(C_n\,\eta)(\sigma) &= \frac{n!}{\phi_0(\sigma_1)\phi_0(\sigma_2)\cdots\phi_0(\sigma_{n-1})} \\ &\quad\times\left(J\left[\frac{1}{(n-1)!}\phi_{0,\,n-1}(C_{n-1}\,\eta)\right]\right)(\sigma) \quad (n>1).\end{aligned}$$

Applying the operation D to this equation, and making use of (4), (VIII), the property (I_D) of J, and the property (R) of $\mathfrak{S}$, we get

$$(D[\phi_0(C_n\eta)])(\sigma) = -\frac{(n-1)(n!)}{\phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-1})}$$
$$\times\left(J\left[\frac{1}{(n-1)!}\phi_{0,\,n-1}(C_{n-1}\eta)\right]\right)(\sigma) + n(C_{n-1}\eta)(\sigma).$$

Applying the operation J to this equation, and making use of property (I_J) and equation (15), we obtain

$$\phi_0(\sigma)(C_n\eta)(\sigma) = -(n-1)(J(C_n\eta))(\sigma) + n(J(C_{n-1}\eta))(\sigma).$$

Transposing the first term on the right-hand side and dividing through by $n\phi_0(\sigma)$, we have finally the identity (14) for $n > 1$. For $n = 1$ it is an obvious consequence of (11).

Before stating the next lemma we need to introduce the following notation:

$$\text{(16)}\qquad \begin{aligned} &\phi_n(\sigma) = \phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-2})\phi(\sigma) \quad (n > 2),\\ &\phi_2(\sigma) = \phi_0(\sigma)\phi(\sigma), \qquad \phi_1(\sigma) = \phi(\sigma). \end{aligned}$$

LEMMA 2. *If* $\lim_\sigma \phi(\sigma)$ *exists and is equal to* a *and* $|\phi(\sigma)| < a_1$ *for every* σ, *then* $\lim_\sigma [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma)$ *will exist and be equal to* a/n *and we shall have*

$$|[\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma)| < \frac{a_1}{n} \qquad (\sigma).$$

Given a positive e, we choose σ'_e so that $a - e/4 < \phi(\sigma) < a + e/4$ for $\sigma > \sigma'_e$. We have

$$\text{(17)}\qquad \begin{aligned} [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma) &= [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma'_e)\\ &\quad + [\phi_{0n}(\sigma)]^{-1}[(J\phi_n)(\sigma) - (J\phi_n)(\sigma'_e)]. \end{aligned}$$

In view of (16), (3), and (VII) we have the relationship

$$\text{(18)}\qquad (J\phi_n)(\sigma) = \left(J\left[\frac{1}{n}(D\phi_{0n})\phi\right]\right)(\sigma).$$

Making use of (18) and postulates M_2 and I_J, we see that the second term on the right-hand side of (17) lies between

$$\frac{1}{n}\left(a - \frac{e}{4}\right)\cdot\left[1 - \frac{\phi_{0n}(\sigma'_e)}{\phi_{0n}(\sigma)}\right] \quad\text{and}\quad \frac{1}{n}\left(a + \frac{e}{4}\right)\cdot\left[1 - \frac{\phi_{0n}(\sigma'_e)}{\phi_{0n}(\sigma)}\right].$$

We see from (IV) that for a proper choice of $\sigma''_e > \sigma'_e$, each of the above expressions differs from a/n by a quantity less in absolute value than $\frac{1}{2}e$ for all $\sigma > \sigma''_e$.

The first term on the right side of (17) is seen from (18) and (M_1) to be less in absolute value than

$$\frac{a_1}{n}\left|\frac{\phi_{0n}(\sigma'_e)}{\phi_{0n}(\sigma)}\right|.$$

It follows from (IV) that we can choose $\sigma'''_e > \sigma'_e$ so as to make this latter expression less in absolute value than $\frac{1}{2}e$ for $\sigma > \sigma'''_e$.

If now we choose for σ_e the greater of σ''_e and σ'''_e, it follows from (17) that for $\sigma > \sigma_e$, $|[\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma) - a/n| < e$. The first statement in our conclusion is therefore established. We may readily infer that the second statement holds also if we note that in view of (18) and postulates M_1 and I_J, we have

$$-\frac{a_1}{n} < [\phi_{0n}(\sigma)]^{-1}(J\phi_n)(\sigma) < \frac{a_1}{n} \qquad (\sigma).$$

Let us set

$$\phi'(\sigma) = \frac{1}{n}\phi(\sigma) + \frac{n-1}{n}\cdot\frac{1}{\phi_0(\sigma)}(J\phi)(\sigma). \tag{19}$$

We shall then prove

LEMMA 3. *If* $\lim_\sigma \phi'(\sigma)$ *exists and is equal to a and* $|\phi'(\sigma)| < a_1$ *for every* σ, *then* $\lim_\sigma \phi(\sigma)$ *will exist and be equal to a and we shall have* $|\phi(\sigma)| < a_2$ *for every* σ.

We define $\phi'_n(\sigma)$ in a manner analogous to that in which $\phi_n(\sigma)$ is defined by (16). Then multiplying (19) by $n\phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-2})$ or by $2\phi_0(\sigma)$ according as $n > 2$ or $n = 2$, and making use of (VII) and (3), we have

$$n\phi'_n(\sigma) = \phi_n(\sigma) + [(D\phi_{0,\,n-1})(\sigma_1)]\cdot[(J\phi)(\sigma)] \qquad (n \geqq 2).$$

Applying the operation J to this equation and making use of (5), (I_D), and (I_J), we obtain

$$\begin{aligned} n(J\phi'_n)(\sigma) &= (J\phi_n)(\sigma) + \phi_{0,\,n-1}(\sigma_1)\cdot(J\phi)(\sigma) - (J\phi_n)(\sigma) \\ &= \phi_{0,\,n-1}(\sigma_1)\cdot(J\phi)(\sigma) \quad (n \geqq 2). \end{aligned}$$

Combining the above equation with (19), we have

$$\phi(\sigma) = n\phi'(\sigma) - \frac{n(n-1)}{\phi_0(\sigma)\phi_0(\sigma_1)\cdots\phi_0(\sigma_{n-1})}(J\phi'_n)(\sigma) \qquad (n \geqq 2).$$

Applying Lemma 2 we see that the second term of the right side of this equation approaches $-(n-1)a$ as a limit and remains finite for all σ. Hence our lemma is proved for the case $n \geqq 2$. For $n = 1$ it is an obvious consequence of (19).

Let us set

$$\left(\left(\frac{n-1}{n} M + \frac{1}{n} E\right)\gamma\right)(\sigma) = (S_n \gamma)(\sigma) \qquad (n).$$

Noting that S_n and M are interchangeable operations, we have, from successive applications of (14),

$$(H_1 \eta)(\sigma) = (M\eta)(\sigma) = (M(C_0 \eta))(\sigma) = (S_1(C_1 \eta))(\sigma),$$

$$(H_2 \eta)(\sigma) = (M(H_1 \eta))(\sigma) = (M(S_1(C_1 \eta)))(\sigma)$$
$$= (S_1(M(C_1 \eta)))(\sigma) = (S_1(S_2(C_2 \eta)))(\sigma),$$

. . . .

. . . .

$$(H_n \eta)(\sigma) = (M(H_{n-1} \eta))(\sigma)$$
$$= (M(S_1(S_2(S_3 \cdots (S_{n-1}(C_{n-1} \eta)) \cdots))))(\sigma)$$
$$= (S_1(S_2(S_3 \cdots (S_{n-1}(M(C_{n-1} \eta))) \cdots)))(\sigma)$$
$$= (S_1(S_2(S_3 \cdots (S_{n-1}(S_n(C_n \eta))) \cdots)))(\sigma).$$

We are now ready to prove the general equivalence theorem:

THEOREM. *If* $\lim_\sigma (C_n \eta)(\sigma)$ *exists and is equal to* a, *then* $\lim_\sigma (H_n \eta)(\sigma)$ *will exist and be equal to* a, *and conversely.*

From the last equation above, property (B), and successive applications of Lemma 2 for the case $n = 1$, we infer that the existence of $\lim_\sigma (C_n \eta)(\sigma) = a$ implies the existence of $\lim_\sigma (H_n \eta)(\sigma) = a$, for every n. From the same equation, property (B), and successive applications of Lemma 3 we draw the converse conclusion. Thus our theorem is established.

UNIVERSITY OF CINCINNATI,
CINCINNATI, OHIO.

ANHARMONIC POLYNOMIAL GENERALIZATIONS OF THE NUMBERS OF BERNOULLI AND EULER*

BY

E. T. BELL

We consider twelve infinite systems of polynomials in z which for $z = 1$ degenerate either to the numbers of Bernoulli or Euler, or to others simply dependent upon these. The first part proceeds from the definition of anharmonic polynomials to the specific twelve systems discussed; the second presents an adaptation of the symbolic calculus of Blissard and Lucas in sufficient detail for rapidly developing a simple isomorphism between the algebra of the polynomials and that of the twelve elliptic functions sn, cn, ns, nc, sc, $\cdots$ of Glaisher, and the third contains a short selection from the simpler algebraic and congruential relations between the polynomials. Incidentally there is pointed out in the second part a new interpretation of Kronecker's work on certain symmetric functions and their connections with Bernoulli's numbers. Owing to the length of the paper the development stops short of the quadratic transformation of the polynomials which corresponds to the transformation of the second order in elliptic functions, but the material given is a necessary foundation for all higher transformations. For the same reason only prime moduli are considered in the congruences, although the case in which the modulus is a power of a prime can be treated in essentially the same way, but at greater length. All references are at the end of the paper.

I. Anharmonic polynomials

1. With each of the substitutions σ' of the cross ratio group on z,

$$1 = (z, z), \quad \alpha' = (z, 1/(1-z)), \quad \beta' = (z, (z-1)/z),$$
$$\gamma' = (z, 1/z), \quad \delta' = (z, 1-z), \quad \epsilon' = (z, z/(z-1)),$$

associate a multiplier σ'' as follows,

$$1, \quad \alpha'' = (z-1)^n, \quad \beta'' = (-z)^n,$$
$$\gamma'' = z^n, \quad \delta'' = (-1)^n, \quad \epsilon'' = (1-z)^n,$$

n being an integer $\geqq 0$. From these define six linear substitutions upon the coefficients $a_0, a_1, \cdots, a_n$, with $a_n \neq 0$, of the polynomial

$$A(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

* Presented to the Society, San Francisco Section, April 9, 1921.

by means of the identities

$$\sigma A(z) \equiv \sigma'' A(\sigma' z) \equiv a_{\sigma 0} + a_{\sigma 1} z + a_{\sigma 2} z^2 + \cdots + a_{\sigma n} z^n,$$

in which $A(\sigma' z)$ denotes the result of applying σ' to $A(z)$. Hence $\sigma A(z)$ is the polynomial derived from $A(z)$ by first operating with σ' and then multiplying the result by σ''. If $a_0 = 0$ the degree of $\sigma A(z)$ is $< n$ for some σ. The six σ's thus obtained are $1, \alpha, \beta, \gamma, \delta, \epsilon$, where α corresponds to α', α'', etc.

2. A set of generating relations for the cross ratio group is

$$\alpha'^3 = \gamma'^2 = (\alpha' \gamma')^2 = 1.$$

We have $\alpha A(z) = (z-1)^n A(\alpha' z)$; whence

$$\alpha^2 A(z) \equiv \alpha(\alpha A(z)) = (z-1)^n [(z/(1-z))^n A(\alpha'^2 z)],$$

$$\alpha^3 A(z) \equiv \alpha(\alpha^2 A(z)) = (z-1)^n [(1/(z-1))^n A(\alpha'^3 z)] = A(z).$$

Hence $\alpha^3 = 1$, and similarly $\gamma^2 = 1$, $(\alpha\gamma)^2 = 1$, so that the σ form a group Γ simply isomorphic to the cross ratio group. For convenience of reference we reproduce its multiplication table, which is to be read in the usual way; thus $\alpha\gamma = \delta$, $\gamma\epsilon = \alpha$, and so on:

	1	α	β	γ	δ	ϵ
α	α	β	1	δ	ϵ	γ
β	β	1	α	ϵ	γ	δ
γ	γ	ϵ	δ	1	β	α
δ	δ	γ	ϵ	α	1	β
ϵ	ϵ	δ	γ	β	α	1

3. Obviously $a_{\gamma r} = a_{1n-r} \equiv a_{n-r}$. Let ρ, σ, τ be substitutions of Γ between which there is the relation $\rho\sigma = \tau$, and write $a_{\rho\sigma r}$ for the function of the coefficients a_r which is derived from the $a_{\sigma r}$ in the same way that $a_{\rho r}$ is from the a_r. Then $a_{\rho\sigma r} = a_{\tau r}$, and the coefficients of the six polynomials are related as shown by the table in § 2. If $\gamma\sigma = \tau$, then $a_{\tau r} = a_{\gamma\sigma r} = a_{\sigma n-r}$. Hence, writing as usual $0! = 1$, $\binom{r}{0} = \binom{0}{0} = 1$, $\binom{r}{s} = r!/(s!(r-s)!)$, we have the following forms of the $a_{\sigma r}$, $r = 0, 1, s, \cdots, n$, which explicitly define the substitutions of Γ:

$$a_{1r} = a_{\gamma n-r} = a_r,$$

$$a_{\alpha r} = a_{\epsilon n-r} = (-1)^{n-r} \sum_{s=0}^{n-r} \binom{n-s}{r} a_s,$$

$$a_{\beta r} = a_{\delta n-r} = (-1)^r \sum_{s=0}^{r} \binom{n-s}{r-s} a_{n-s},$$

the $a_{\sigma r}$ being given directly by the $\sigma A(z)$ for $\sigma = 1, \alpha, \beta$.

4. The set $A(z)$, $\alpha A(z)$, $\beta A(z)$, $\gamma A(z)$, $\delta A(z)$, $\epsilon A(z)$ is called anharmonic of degree n, n being the highest degree of any polynomial in the set. Any symmetric function of the polynomials $A(z)$, $\sigma A(z)$ is an invariant of the subgroup 1, σ ($\sigma = \gamma, \delta, \epsilon$); any symmetric function of $A(z)$, $\alpha A(z)$, $\beta A(z)$ is an invariant of the subgroup 1, α, β; and any symmetric function of $\sigma A(z)$ ($\sigma = 1, \alpha, \beta, \gamma, \delta, \epsilon$) is an invariant of Γ.

5. When $A(z)$ is reciprocal, $\gamma A(z) = A(z)$. Hence from the multiplication table $\delta A(z) = \alpha A(z)$, $\epsilon A(z) = \beta A(z)$, and in this case the anharmonic set reduces to $A(z)$, $\alpha A(z)$, $\beta A(z)$. We shall call such a set cyclic, and henceforth reserve the term anharmonic for sets that are not cyclic or, what is equivalent, contain no reciprocal polynomial. Any symmetric function of all the members of a cyclic set is invariant for $1, \alpha, \beta$.

6. The σ are determined by the coefficients a_r as in § 3 and are functions of the degree n of the set. When necessary to designate the polynomial fixing the σ we shall speak of the associated σ as the cross ratio substitution for that polynomial. To indicate that $A(z)$ belongs to a set of degree n we write $A_{(n)}(z)$, enclosing the n in ( ) to distinguish it from the rank defined in § 7. When it is a question of relations between the coefficients of polynomials of different degrees (§ 15) the a_r may be given double suffixes. Thus

$$A(z) \equiv A_{(n)}(z) \equiv \sum_{r=0}^{n} a_r z^r \equiv \sum_{r=0}^{n} a_{nr} z^r$$

are merely different notations for one polynomial, and likewise for

$$\sigma A(z) \equiv \sigma A_{(n)}(z) \equiv \sum_{r=0}^{n} a_{\sigma r} z^r \equiv \sum_{r=0}^{n} a_{\sigma n r} z^r .$$

Where there can be no confusion we shall use the simpler forms.

7. Let x, z be independent variables, and put $\sqrt{z} = k$, $\sqrt{1-z} = k'$, $i = \sqrt{-1}$. Let t, v be functions of x, k (or of x, z) such that

$$t(x, -k) = t(x, k) = -t(-x, k),$$

$$v(x, -k) = v(x, k) = v(-x, k),$$

and assume (cf. § 12) that these functions can be expanded in absolutely convergent power series in x of the form

$$t(x, k) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} T_{(n)}(z), \qquad v(x, k) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} V_{(n)}(z).$$

We now define $T_{(n)}(z)$ arising from an odd function generator $t(x, k)$ to be an odd polynomial of rank $2n+1$, and $V_{(n)}(z)$ generated by an even function $v(x, k)$ to be an even polynomial of rank $2n$, and write

$$T_{(n)}(z) \equiv T_{2n+1}(z), \qquad V_{(n)}(z) \equiv V_{2n}(z),$$

so that the degree of a set containing a given polynomial is the greatest integer in half the rank. This change in notation is essential for the further development, as without it the application of the symbolic calculus of Blissard and Lucas is impracticable. Exhibiting the ranks rather than the degrees we have

$$t(x,k)=\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!}T_{2n+1}(z)\equiv \sin T(z)x,$$

$$v(x,k)=\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n)!}V_{2n}(z)\equiv \cos V(z)x,$$

the trigonometric forms being the purely symbolic equivalents of the series.

The principles of the symbolic method, which we shall use freely in the sequel, having been fully treated by Blissard, Lucas and Glaisher, in the works cited, we need not recall them here, except to emphasize the caution that in all operations with symbolic powers zero exponents must be included. Thus the first term of $\cos V(z)x$ is $V_0(z)$, not unity; the symbolic binomial $(p+q)^2=p^2q^0+2p^1q^1+p^0q^2\equiv p_2q_0+2p_1q_1+p_0q_2$, not $p_2+2p_1q_1+q_2$; and $(p-p)^2=2(p_2p_0-p_1^2)$, obtained from $(p-q)^2$ by putting $q=p$ in the final form of the latter.

8. That σ is one of the cross ratio substitutions for $A_n(z)$ may be indicated by writing σ_n, but for simplicity we shall put $\sigma_n A_n(z)\equiv \sigma A_n(z)$. Analogously to the sn, cn, ns, $\cdots$ notation for elliptic functions we denote each of the six polynomials $1A_n(z)\equiv A_n(z)$, $\alpha A_n(z)$, $\beta A_n(z)$, $\cdots$, $\epsilon A_n(z)$ by a double letter symbol $1A\equiv A$, αA, βA, $\cdots$, ϵA, and in any such σA regard the σ, A as inseparable. In Blissard's method $A_n(z)$ is written $A^n(z)$, or A^n when z is understood, the exponent being purely symbolic, and A is called an umbra. Similarly we now have $(\sigma A)^n\equiv\sigma A_n$, since σA is one symbol, and our umbræ are double-letter symbols σA. It is important to note once for all that $(\sigma A)^n$ is not $\sigma^n A_n$ in which σ^n has the usual meaning as a power of a substitution.

9. We require the operations transforming the generators of T, V into those of σT, σV respectively. Let $f(x,k)$ denote an arbitrary function of x, k, and Ω_i an operator which applied to f transforms it as follows:

$$\Omega_i f(x,k)=\phi_i(k)\,f(x\psi_i(k),\chi_i(k)).$$

We regard Ω_i as a tripartite operator,

$$\Omega_i\equiv\{\phi_i(k),\psi_i(k),\chi_i(k)\},$$

which replaces k in $f(x,k)$ by $\chi_i(k)$, x by $x\psi_i(k)$, and multiplies the function thus transformed by $\phi_i(k)$. Hence the product $\Omega_j\Omega_i$ in which Ω_i is applied first is

$$\{\phi_j(k)\phi_i(\chi_j(k)),\quad \psi_j(k)\psi_i(\chi_j(k)),\quad \chi_i(\chi_j(k))\};$$

and if $f(x, k)$ is odd in x, even in k,

$$\{\phi, \psi, \chi\} = \{-\phi, -\psi, \chi\} = \{\phi, \psi, -\chi\};$$

while if $f(x, k)$ is even in both x and k,

$$\{\phi, \psi, \chi\} = \{\phi, -\psi, \chi\} = \{\phi, \psi, -\chi\}.$$

Let σ', σ denote any cross ratio substitutions for $T_{2n+1}(z)$, $V_{2n}(z)$ respectively, so that symbolically

$$\sin \sigma' T(z) x \equiv \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \sigma' T_{2n+1}(z),$$

$$\cos \sigma V(z) x \equiv \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \sigma V_{2n}(z);$$

and designate by s', s operators such that

$$s' t(x, k) = \sin \sigma' T(z) x, \qquad sv(x, k) = \cos \sigma V(z) x.$$

Then by inspection of the series, the Greek and Latin letters (α, a), (α', a'), $\cdots$, (σ, s), (σ', s') corresponding, we have

$$\begin{aligned}
1 &= \{1, 1, k\}, & 1 &= \{1, 1, k\}, \\
a' &= \{-i/k', ik', 1/k'\}, & a &= \{1, ik', 1/k'\}, \\
b' &= \{-i/k, ik, ik'/k\}, & b &= \{1, ik, ik'/k\}, \\
c' &= \{1/k, k, 1/k\}, & c &= \{1, k, 1/k\}, \\
d' &= \{-i, i, k'\}, & d &= \{1, i, k'\}, \\
e' &= \{1/k', k', ik/k'\}; & e &= \{1, k', ik/k'\}.
\end{aligned}$$

From the definitions the s' form a group G' which is simply isomorphic to Γ and whose multiplication table is obtained from that in § 2 upon replacing each Greek letter by its accented Latin correspondent; the multiplication table for the group G of operators s is obtained from G' by suppressing accents. The explicit forms of the operators given above can be verified easily on combining them according to the formulas developed for the Ω_i, noting that the s' refer to an odd, and the s to an even function.

10. Thus far s' has been applied only to the odd $t(x, k)$, and s to the even $v(x, k)$. The following cases are of equal importance, and they may be seen at once from the definitions. If

$$s' \equiv \{\phi'(k), \psi'(k), \chi'(k)\}, \qquad \text{then} \qquad s = \{1, \psi'(k), \chi'(k)\};$$

$$s' v(x, k) = \phi'(k) \cdot sv(x, k), \qquad st(x, k) = \frac{1}{\phi'(k)} \cdot s' t(x, k).$$

11. We now specialize the even V in each of two ways and the odd T in one way, getting in all for each integral degree $n \geqq 0$ a system of twelve polynomials distributed into two cyclic sets and one anharmonic. Indicating that the modulus of the elliptic functions is k by writing $\operatorname{sn}(x, k)$, etc., we define the fundamental polynomials S, C, P and their anharmonic transforms $\sigma' S$, σC, σP by their symbolic generators,

$$\begin{aligned} \operatorname{sn}(x, k) &\equiv \sin S(z)x, & s' \operatorname{sn}(x, k) &\equiv \sin \sigma' S(z)x, \\ \operatorname{cn}(x, k) &\equiv \cos C(z)x, & s \operatorname{cn}(x, k) &\equiv \cos \sigma C(z)x, \\ x \operatorname{ns}(x, k) &\equiv \cos P(z)x; & sx \operatorname{ns}(x, k) &\equiv \cos \sigma P(z)x. \end{aligned}$$

From the elements of elliptic functions the coefficients in $S_{2n+1}(z)$, $C_{2n}(z)$ are positive integers, $C_{2n}(z)$ is of degree $n-1$ in z, and S_{2n+1} is a reciprocal polynomial of degree n in z. Hence $P_{2n}(z)$ is a reciprocal polynomial, since $x \operatorname{ns} x = x/\operatorname{sn} x$. Therefore, omitting ranks from the notation, the coefficients in each of the sets $\sigma' S(z)$, $\sigma C(z)$ are integers, those in $\sigma P(z)$ are rational but not integral, and of the three sets $\sigma' S(z)$, $\sigma P(z)$ are cyclic, $\sigma C(z)$ is anharmonic.

12. If for all values of z whose absolute value does not exceed a constant different from zero the algebraic relation $F(z) = 0$ holds, then it is easily seen that $F(z') = 0$ is an identity in arbitrary z'. Hence in all polynomial formulas we consider z arbitrary, it being understood that when $F(z)$ is regarded as a coefficient in an infinite series z is such as to render the series absolutely convergent (all the series discussed have radius of convergence > 0), but that in all other connections z is arbitrary.

13. For convenience of comparison with Glaisher's grouping of the elliptic functions into triads we give the complete set of generators:

$$\begin{aligned} \operatorname{sn}(x, k) &= \sin S(z)x, & x \operatorname{ns}(x, k) &= \cos P(z)x, \\ \operatorname{cn}(x, k) &= \cos C(z)x, & x \operatorname{cs}(x, k) &= \cos \alpha P(z)x, \\ \operatorname{dn}(x, k) &= \cos \gamma C(z)x; & x \operatorname{ds}(x, k) &= \cos \beta P(z)x; \\ \operatorname{dc}(x, k) &= \cos \alpha C(z)x, & \operatorname{cd}(x, k) &= \cos \epsilon C(z)x, \\ \operatorname{nc}(x, k) &= \cos \delta C(z)x, & \operatorname{sd}(x, k) &= \sin \beta S(z)x, \\ \operatorname{sc}(x, k) &= \sin \alpha S(z)x; & \operatorname{nd}(x, k) &= \cos \beta C(z)x. \end{aligned}$$

Glaisher points out that of the four triads the second is the most symmetrical, and that it should be taken as a basis for the jacobian elliptic functions rather than the traditional first. From the present point of view the second triad frequeptly appears to be wholly anomalous: while the sets $\sigma' S(z)$, $\sigma C(z)$ in many significant ways can be regarded as forming one complete system, the

set $\sigma P(z)$ stands apart in the majority of relations of a specific type. We see in a moment that $\sigma P(z)$ is related to the Bernoulli numbers, $\sigma' S(z)$, $\sigma C(z)$ to those of Genocchi and Euler respectively. Hence we have another instance of that superficial similarity and radical difference between the numbers of Bernoulli and Euler which has often been remarked.

14. When $z = 1$ and hence $k = 1$ the elliptic functions degenerate to circular functions of the gudermannian of x. From the values of these as given by Cayley, p. 59, and the symbolic generators for the B, E, G, R (numbers of Bernoulli, Euler, Genocchi and Lucas) in Lucas 5, p. 262, we find at once

$$\begin{aligned} x \text{ ns}(x, 1) &= ix \cot ix = \cos 2Bix, \\ \text{cn}(x, 1) &= \sec ix = \cos Eix, \\ \text{dn}(x, 1) &= \sec ix = \cos Eix, \\ 2x \text{ sn}(x, 1) &= -2ix \tan ix = -\cos 2Gix, \\ x \text{ ds}(x, 1) &= ix \operatorname{cosec} ix = 2\cos Rix, \\ x \text{ cs}(x, 1) &= ix \operatorname{cosec} ix = 2\cos Rix. \end{aligned}$$

The even suffix notation of Lucas is used for all B, E, G, R, and the last two are defined by

$$G_{2n} = 2(1 - 2^{2n})B_{2n}, \qquad R_{2n} = (1 - 2^{2n-1})B_{2n}.$$

We require a fifth system of numbers H_{2n+1}, the so-called tangent coefficients, shown presently to be integers > 0,

$$H_{2n+1} = \frac{(-1)^{n+1}}{n+1} 2^{2n} G_{2n+2} = \frac{(-1)^n}{n+1} 2^{2n+1}(2^{2n+2} - 1)B_{2n+2}.$$

Comparing coefficients of like powers of x in the generators above with those in § 13 we have for $n \geqq 0$, $r > 0$ (note that in each case the degenerate form is expressed as a function of the rank $2n$ or $2n + 1$),

$$\begin{aligned} &S_{2n+1}(1) = H_{2n+1}, && P_{2n}(1) = (-1)^n 2^{2n} B_{2n} = (2i)^{2n} B_{2n}, \\ &\alpha S_{2n+1}(1) = (-1)^n = -i \cdot i^{2n+1}, && \alpha P_{2n}(1) = 2(-1)^n R_{2n} = 2i^{2n} R_{2n}, \\ &\beta S_{2n+1}(1) = (-1)^n = -i \cdot i^{2n+1}; && \beta P_{2n}(1) = 2(-1)^n R_{2n} = 2i^{2n} R_{2n}; \\ &C_{2n}(1) = (-1)^n E_{2n} = i^{2n} E_{2n}, && \gamma C_{2n}(1) = (-1)^n E_{2n} = i^{2n} E_{2n}, \\ &\alpha C_0(1) = 1, \alpha C_{2r}(1) = 0, && \delta C_{2n}(1) = (-1)^n = i^{2n}, \\ &\beta C_{2n}(1) = (-1)^n = i^{2n}; && \epsilon C_0(1) = 1, \quad \epsilon C_{2r}(1) = 0. \end{aligned}$$

The coefficients in $S_{2n+1}(z)$, $C_{2n}(z)$ being integers > 0, so also are H_{2n+1}, $(-1)^n E_{2n}$.

15. An identity between some or all of the twelve elliptic functions implies and is implied by the identity between polynomials which is obtained upon equating coefficients of like powers of the argument x, and as degenerate cases for $z = 1$ of these identities we have relations between the numbers B, E, H, R. The elliptic identities may conveniently be segregated into classes according to the groups G, G' of § 9. All the identities arising from a given one by successive applications of the operations of G belong to the first class, all those similarly derived by means of G' belong to the second, and there are subsidiary classes corresponding to the cyclic subgroups of orders $3, 2$. Polynomial identities derived from elliptic identities of a given class belong to one class. Elliptic identities are further subdivided into types according to the degree of the identity in sn, cn, $\cdots$, and the derived polynomial relations are similarly subdivided, the degenerate cases being included. By means of the table in § 2 and those next given, together with the formulas of § 10, we write down immediately from any elliptic identity all others of the same type, and hence on replacing the elliptic functions by their symbolic trigonometric equivalents from § 13 we at once infer all the polynomial relations of one type. It is clear that once the elliptic identity is given the rest of the process demands very little labor. In the next part we develop the symbolic method proper to the subject, and this still further reduces the algebra.

Consider any one of the polynomial relations. This is an identity in z and hence, z being arbitrary, it is equivalent to a set of identical relations between the coefficients of the several polynomials. The coefficients of $S_{2n+1}(z)$ are $n + 1$ integers > 0 into which H_{2n+1} is partitioned; those of $C_{2n}(z)$ are n integers > 0 into which $(-1)^n E_{2n}$ is partitioned, and those of $P_{2n}(z)$ are a rational but not integral partition of $(-1)^n 2^{2n} B_{2n}$. From the explicit values of the transformed coefficients in § 3 the coefficients of the transforms of the fundamental polynomials S, C, P are known in terms of the foregoing partitions, and hence the polynomial relation is equivalent to a theorem concerning partitions of a certain kind of the numbers B, E, H of various ranks. Similar remarks apply to congruences between such of the polynomials as take integral values when z is an integer. This aspect is not further elaborated here, as its complete discussion presupposes a knowledge of the arithmetical form of the coefficients in $C_{2n}(z)$, $S_{2n+1}(z)$.* It is not difficult to give implicit arithmetical definitions of these coefficients, but this is not what is required.

* The well known method of Hermite (10, pp. 265, 269) for calculating $C_{2n}(z)$, $S_{2n+1}(z)$ does not give the required information, as recognized by Hermite himself (12, p. 237). His second solution (ibid.), as pointed out by the editors of his works, unfortunately is erroneous, and even if it were correct it is difficult to see what the general coefficient would be from the forms of those given. Writers on elliptic functions seem to have overlooked Hermite's remarks in the second citation, and to have assumed that his first paper is sufficient.

16. From the multiplication tables of G, G' combined with the results of the linear transformations of $\text{sn}(x,k)$, $\text{cn}(x,k)$ (as given for example by Glaisher (6, p. 120)) we have the following for the generators of the $\sigma' S$, σP:

	sn	sc	sd
1	sn	sc	sd
a'	sc	sd	sn
b'	sd	sn	sc

	x ns x	x cs x	x ds x
1	x ns x	x cs x	x ds x
a'	x cs x	x ds x	x ns x
b'	x ds x	x ns x	x cs x

the argument being x and the modulus k. Since each set is cyclic the transformations c', d', e' are respectively identical with 1, a', b'. Corresponding to § 5 we have: any symmetric function of all the members of either of these cyclic sets is an invariant of 1, a', b'.

For the anharmonic set the table is

	cn	dc	nd	dn	nc	cd
1	cn	dc	nd	dn	nc	cd
a	dc	nd	cn	nc	cd	dn
b	nd	cn	dc	cd	dn	nc
c	dn	cd	nc	cn	nd	dc
d	nc	dn	cd	dc	cn	nd
e	cd	nc	dn	nd	dc	cn

The modulus and argument in each case are k, x. Any symmetric function of the members of the following pairs (cn, dn), (dc, dn), (dn, nd) is an invariant of the group $(1, c)$, $(1, d)$ or $(1, e)$ respectively; any symmetric function of dn, nc, cd is an invariant of $(1, a, b)$, and any symmetric function of cn, dc, nd, dn, nc, cd is an invariant of G.

17. One example will suffice to show how all the relations of a given type may be written down from one of them. When k or k' occurs as a factor it is replaced in the final result by its z-equivalent. The modulus being k, consider $\text{sn}^2 x + \text{cn}^2 x = 1$. From this, since $\phi(k) = 1$ for each s,

$$(s\,\text{sn}\,x)^2 + (s\,\text{cn}\,x)^2 = 1\,.$$

Putting $s = a$ we have, by § 10 and the form of a' in § 9,

$$(ik'\,a'\,\text{sn}\,x)^2 + (a\,\text{cn}\,x)^2 = 1;$$

and hence from the tables in § 16, $-k'^2\,\text{sc}^2 x + \text{dc}^2 x = 1$. To illustrate useful processes we shall also consider in detail the effects of operating with s'. Let $s' \equiv \{\phi', \psi', \chi'\}$. Then

$$s'[\text{sn}^2 x + \text{cn}^2 x] = s'\,1 = \phi'(k);$$

$$\phi'(k)[\text{sn}^2(x\psi'(k), \chi'(k)) + \text{cn}^2(x\psi'(k), \chi'(k))] = \phi'(k);$$

$$\text{sn}^2\,(x\psi'(k),\,\chi'(k)) + \text{cn}^2\,(x\psi'(k),\,\chi'(k)) = 1;$$

$$\frac{1}{\phi'^2(k)}\,(s'\,\text{sn}\,(x,k))^2 + (s\,\text{cn}\,(x,k))^2 = 1,$$

and hence for $s' = a'$, $-\,k'^2\,\text{sc}^2\,x + \text{dc}^2\,x = 1$.

With a little practice the transform can be written down by inspection in any case directly from §§ 9, 10, 16. Thus the complete set here is

(1), $\qquad \text{cn}^2\,x + \text{sn}^2\,x = 1,$

$(a \text{ or } a')$, $\qquad \text{dc}^2\,x - k'^2\,\text{sc}^2\,x = 1,$

$(b \text{ or } b')$, $\qquad \text{nd}^2\,x - k^2\,\text{sd}^2\,x = 1,$

$(c \text{ or } c')$, $\qquad \text{dn}^2\,x + k^2\,\text{sn}^2\,x = 1,$

$(d \text{ or } d')$, $\qquad \text{nc}^2\,x - \text{sc}^2\,x = 1,$

$(e \text{ or } e')$, $\qquad \text{cd}^2\,x + k'^2\,\text{sd}^2\,x = 1;$

whence the final forms

$$\cos^2 C(z)x + \sin^2 S(z)x = 1,$$

$$\cos^2 \alpha C(z)x - (1-z)\sin^2 \alpha S(z)x = 1,$$

$$\cos^2 \beta C(z)x - z\sin^2 \beta S(z)x = 1,$$

$$\cos^2 \gamma C(z)x + z\sin^2 S(z)x = 1,$$

$$\cos^2 \delta C(z)x - \sin^2 \alpha S(z)x = 1,$$

$$\cos^2 \epsilon C(z)x + (1-z)\sin^2 \beta S(z)x = 1.$$

Simple rules may be devised for writing down the appropriate multipliers such as z, $1 - z$ above, from the forms of the s', s and § 10, for the several terms of any symbolic identity when the argument z of the polynomials is transformed by the substitutions of Γ. As these present no difficulty we omit them.

II. Isomorphism with elliptic functions

18. To find the polynomial relation equivalent to a given elliptic identity we evidently must consider the properties of products of t symbolic sines and r symbolic cosines in the cases $r,\, t > 0$; $r > 0,\, t = 0$; $r = 0,\, t > 0$. When several symbolic factors of a product are identical we proceed as in the following example. Let λ, μ denote umbræ (§ 8), and suppose the coefficient of x^n in $\cos^2 \lambda x$ is required. We write $\cos^2 \lambda x \equiv \cos \lambda x \cos \mu x$, find the coefficient of x^n in $\cos \lambda x \cos \mu x$ by actual multiplication of the series for $\cos \lambda x$, $\cos \mu x$ or otherwise (§ 24), and in the result, *after* each exponent of λ, μ has been degraded

to a suffix, replace μ by λ. Once more we emphasize that in all expansions zero exponents must be included. Thus $\lambda, \mu, \cdots, \nu$ being umbræ,

$$(\lambda + \mu + \cdots + \nu)^n = \sum \frac{n!}{r!\,p!\cdots t!} \lambda^r \mu^p \cdots \nu^t = \sum \frac{n!}{r!\,p!\cdots t!} \lambda_r \mu_p \cdots \nu_t,$$

the $\sum$ extending to all $r \geqq 0, p \geqq 0, \cdots, t \geqq 0$ such that

$$r + p + \cdots + t = n.$$

As always, $0! = 1$, and after the completion of all formal operations (multiplications, divisions, etc.) $\lambda^r \mu^p \cdots \nu^t$ is to be replaced by $\lambda_r \mu_p \cdots \nu_t$. By convention, for *any* particular choice of the signs,

$$(\pm \lambda^l \pm \mu^m \pm \cdots \pm \nu^n)^0 = \lambda^0 \mu^0 \cdots \nu^0 \equiv \lambda_0 \mu_0 \cdots \nu_0.$$

19. All letters $\lambda, \cdots, \pi'$ denote umbræ, and the sets $\Lambda, \Lambda', Z, Z' \equiv (\lambda, \mu, \cdots, \rho), (\lambda', \mu', \cdots, \tau'), (\zeta, \xi, \cdots, \phi), (\zeta', \xi', \cdots, \pi')$ contain respectively r, t, f, p letters. For $n \geqq 0$ write

$$(\Lambda | \Lambda')^n \equiv \sum \pm (\pm \lambda \pm \mu \pm \cdots \pm \rho \pm \lambda' \pm \mu' \pm \cdots \pm \tau')^n,$$

in which all the exponents are symbolic, the summation extends to the 2^{r+t} possible combinations of signs within the parentheses, and the outer sign in each case is the product of the signs of $\lambda', \mu', \cdots, \tau'$. The important special cases $t = 0, r = 0$ give respectively

$$(\Lambda | \,)^n = \sum (\pm \lambda \pm \mu \pm \cdots \pm \rho)^n,$$
$$(\,| \Lambda')^n = \sum \pm (\pm \lambda' \pm \mu' \pm \cdots \pm \tau')^n.$$

The *umbral factors* of $(\Lambda | \Lambda')^n$ are by definition $\lambda, \mu, \cdots, \rho, \lambda', \mu', \cdots, \tau'$, and similarly for the others. If in these the letters be interpreted as ordinary quantities and the exponents as algebraic we have precisely the symmetric functions considered by Kronecker, cf. § 26.

By definition the respective types of $(\Lambda | \Lambda')^n$, $(\Lambda | \,)^n$, $(\,| \Lambda')^n$ are $(r | t)$, $(r | 0)$, $(0 | t)$, and the weight of each is n, = the sum of the suffixes in the final non-symbolic forms of each. Considered as functions of $\lambda_j, \cdots, \tau'_k$ these final forms are homogeneous of degrees $r + t, r, t$ respectively, the degree in any case being equal to the total number of letters in the bar function $(\Lambda | \Lambda')^n$, etc. The properties of these functions $(\Lambda | \Lambda')^n, \cdots$ are immediate from the expressions in § 20 for their symbolic generators. From the principles of the symbolic calculus as developed by Blissard and Lucas it is evident that products, etc., of symbolic sines and cosines can be combined formally according to the rules of trigonometry, and that the coefficients of like powers of x in the several transforms thus effected of any identity are equal. By starting from the elliptic functions of a pure imaginary argument the development can be

carried out isomorphically to the theory of the hyperbolic instead of the circular functions. This in some respects is preferable, but being committed to the other by Lucas' trigonometric generators for B, E, G, R in § 14 we shall not follow it.

20. Replacing each symbolic sine or cosine by its exponential equivalent we see immediately that

$$2^{r+t}\cos\lambda x\cos\mu x\cdots\cos\rho x\sin\lambda' x\sin\mu' x\cdots\sin\tau' x$$

$$= (-1)^{\frac{t-1}{2}}\sin(\Lambda|\Lambda')x \quad\text{or}\quad (-1)^{\frac{t}{2}}\cos(\Lambda|\Lambda')x$$

according as t is odd or even;

$$2^r\cos\lambda x\cos\mu x\cdots\cos\rho x = \cos(\Lambda|\,)x;$$

and according as t is odd or even,

$$2^t\sin\lambda' x\sin\mu' x\cdots\sin\tau' x = (-1)^{\frac{t-1}{2}}\sin(\,|\Lambda')x \quad\text{or}\quad (-1)^{\frac{t}{2}}\cos(\Lambda'|\,)x,$$

all of which are included in the first. Hence $(\Lambda|\Lambda')^n$ is an even function of each of the r letters in Λ, and an odd function of each of the t letters in Λ'; or, $(\Lambda|\Lambda')^n$ of type $(r|t)$ is r-fold even and t-fold odd. Moreover $(\Lambda|\Lambda')^n$ is symmetric in the letters in Λ, also in those in Λ'. Again, from the generators, $(\Lambda|\Lambda')^{2n} = 0$ if t is odd, while $(\Lambda|\Lambda')^{2n+1} = 0$ if t is even. The corresponding statements for $(\Lambda|\,)^n$, $(\,|\Lambda')^n$ are included as special cases.

21. In $(\Lambda|\Lambda')^n$ we are concerned with (1) the umbræ $\lambda, \cdots, \tau'$; (2) the type $(r|t)$; (3) the weight n. Each of these has a species of addition theorem. Denote by (Λ, Z) the set consisting of all the letters in Λ together with all those in Z, and similarly for (Λ', Z'). Then $((\Lambda, Z)|(\Lambda', Z'))^n$ is of type $((r+f)|(t+p))$, and is symmetric in Λ, Z, also in Λ', Z'. The addition theorems with respect to types are given by the following, all of which are obvious on remarking that the generator of $((\Lambda, Z)|(\Lambda', Z'))^n$ is the product of the generators of $(\Lambda|\Lambda')^n$, $(Z|Z')^n$, and reapplying the several cases of the generators in § 20 to these products before equating coefficients of like powers of x. Let $(n, t, p) \equiv (n', t', p') \bmod 2$; then

(n', t', p')	$4((\Lambda, Z)\|(\Lambda', Z'))^n$	$((r+f)\|(t+p))$
$(0, 0, 0)$	$((\Lambda\|\Lambda'), (Z\|Z')\|)^n$	$((r\|t), (f\|p)\|)$
$(0, 1, 1)$	$(\|(\Lambda\|\Lambda'), (Z\|Z'))^n$	$(\|(r\|t), (f\|p))$
$(1, 0, 1)$	$((\Lambda\|\Lambda')\|(Z\|Z'))^n$	$((r\|t)\|(f\|p))$
$(1, 1, 0)$	$((Z\|Z')\|(\Lambda\|\Lambda'))^n$	$((f\|p)\|(r\|t))$.

For example, when n, t, p are all even, a function of type $((r+f)|(t+p))$ is expressible linearly in terms of functions of types $(r|t)$, $(f|p)$ in the manner

shown in the second column. The calculation of such a function is thus reduced to a series of symbolic multiplications (and subsequent additions) of functions whose types are $(r|t)$, $(f|p)$.

22. Omitting the addition theorems for the umbræ we restrict the discussion of those for the weights to the particularly important cases of two and three letters. The general case of n umbræ is treated similarly. In all that follows multiplications indicated by dots are purely symbolic. These are performed analogously to algebraic multiplications but by the addition of suffixes instead of exponents. Thus

$$(\lambda_r + \mu_r)\cdot(\lambda_t + \mu_t) = \lambda_r\cdot\lambda_t + \lambda_r\cdot\mu_t + \mu_r\cdot\lambda_t + \mu_r\cdot\mu_t$$
$$= \lambda_{r+t} + \lambda_r\,\mu_t + \mu_r\,\lambda_t + \mu_{r+t};$$

$\lambda_2\cdot\lambda_3 = \lambda_5$; $\lambda_2\cdot\lambda_2 = \lambda_4$, not λ_2^2. Algebraic multiplication is a special case of this.

Taking first the case of two umbræ λ, μ we define auxiliary functions $\{\,\}$ by

$$2\{\lambda, \mu|\}^r \equiv \phi^r + \psi^r, \qquad 2\{|\lambda, \mu\}^r \equiv \phi^r - \psi^r,$$
$$\phi \equiv \lambda + \mu, \qquad \psi \equiv \lambda - \mu, \qquad r \geqq 0,$$

and put $L^r \equiv L_r = \{\lambda, \mu|\}^r$, $M^r \equiv M_r = \{|\lambda, \mu\}^r$. Then

$$4L^{2r} = (\lambda, \mu|)^{2r}, \qquad 4M^{2r} = (|\lambda, \mu)^{2r}, \qquad 4M^{2r+1} = (\lambda|\mu)^{2r+1}.$$

We have $\phi^r = L^r + M^r$, $\psi^r = L^r - M^r$, and hence for $t \geqq 0$,

$$\phi^{r+t} = (L_r + M_r)\cdot(L_t + M_t), \qquad \psi^{r+t} = (L_r - M_r)\cdot(L_t - M_t).$$

On the other hand

$$2L^{r+t} = \phi^{r+t} + \psi^{r+t}, \qquad 2M^{r+t} = \phi^{r+t} - \psi^{r+t}.$$

Whence, substituting for ϕ^{r+t}, ψ^{r+t} and degrading exponents, we have the addition theorems

$$L_{r+t} = L_r\cdot L_t + M_r\cdot M_t, \qquad M_{r+t} = L_r\cdot M_t + M_r\cdot L_t.$$

For example $L_1 = \lambda_1\,\mu_0$, $L_2 = \lambda_2\,\mu_0 + \mu_2\,\lambda_0$, $M_1 = \lambda_0\,\mu_1$, $M_2 = 2\lambda_1\,\mu_1$; hence $L_1\cdot L_2 = \lambda_3\,\mu_0 + \lambda_1\,\mu_2$, $M_1\cdot M_2 = 2\lambda_1\,\mu_2$. Again

$$L_3 = \tfrac{1}{2}[(\lambda + \mu)^3 + (\lambda - \mu)^3] = \lambda_3\,\mu_0 + 3\lambda_1\,\mu_2,$$

so that $L_3 = L_1\cdot L_2 + M_1\cdot M_2$ as required by the first of the theorems.

23. The case* of three umbræ λ, μ, ν is treated in the same way, and we need

* If λ, μ, ν be interpreted as ordinaries and exponents as algebraic, the special cases of the addition theorems of the first kind (in § 21) for the functions $\{\lambda, \mu, \nu|\}^r$, etc., which are of the same form as those for $(\lambda, \mu, \nu|)^r$, etc., have interesting consequences when r is prime for Fermat's quotients. Cf. Bachmann, Journal für Mathematik, vol. 142 (1913), pp. 41–50; Dickson, *History of the Theory of Numbers*, vol. I, p. 111.

give only the results. There can be no confusion between $\alpha, \beta, \gamma, \delta$ here and the substitutions of Γ, nor between B and the Bernoulli numbers. Write

$$\begin{aligned} \alpha &= \lambda + \mu + \nu, & 4A^r &\equiv 4\{\lambda, \mu, \nu|\}^r \equiv \alpha^r + \beta^r + \gamma^r + \delta^r, \\ \beta &= \lambda - \mu + \nu, & 4B^r &\equiv 4\{\lambda, \mu|\nu\}^r \equiv \alpha^r + \beta^r - \gamma^r - \delta^r, \\ \gamma &= \lambda + \mu - \nu, & 4C^r &\equiv 4\{|\lambda, \mu, \nu\}^r \equiv \alpha^r - \beta^r - \gamma^r + \delta^r, \\ \delta &= \lambda - \mu - \nu, & 4D^r &\equiv 4\{\nu|\lambda, \mu\}^r \equiv \alpha^r - \beta^r + \gamma^r - \delta^r, \end{aligned}$$

and put $A^r, B^r, C^r, D^r \equiv A_r, B_r, C_r, D_r$. We have

$$8A^{2r} = (\lambda, \mu, \nu|)^{2r}, \qquad 8B^{2r+1} = (\lambda, \mu|\nu)^{2r+1},$$
$$8C^{2r+1} = (|\lambda, \mu, \nu)^{2r+1}, \quad 8D^{2r} = (\nu|\lambda, \mu)^{2r},$$

and the addition theorem for $r, t \geqq 0$,

$$\begin{aligned} A_{r+t} &= A_r \cdot A_t + B_r \cdot B_t + C_r \cdot C_t + D_r \cdot D_t, \\ B_{r+t} &= A_r \cdot B_t + B_r \cdot A_t + C_r \cdot D_t + D_r \cdot C_t, \\ C_{r+t} &= A_r \cdot C_t + B_r \cdot D_t + C_r \cdot A_t + D_r \cdot B_t, \\ D_{r+t} &= A_r \cdot D_t + B_r \cdot C_t + C_r \cdot B_t + D_r \cdot A_t. \end{aligned}$$

There also are multiplication theorems for the calculation of A_{nr}, etc., but we shall omit these.

24. As they are frequently useful we write down the $\{\}$ equivalents for 2 and 3 letters of the generators in § 20:

$$\begin{aligned} \cos \lambda x \cos \mu x &= \cos \{\lambda, \mu|\}x = \cos \{\mu, \lambda|\}x, \\ \sin \lambda x \sin \mu x &= -\cos \{|\lambda, \mu\}x = -\cos \{|\mu, \lambda\}x, \\ \cos \lambda x \sin \mu x &= \sin \{\lambda|\mu\}x = \sin \{\mu, \lambda|\}x, \\ \sin \lambda x \cos \mu x &= \sin \{\mu|\lambda\}x = \sin \{\lambda, \mu|\}x; \\ \cos \lambda x \cos \mu x \cos \nu x &= \cos \{\lambda, \mu, \nu|\}x = \cos \{\{\lambda, \mu|\}, \nu|\}x, \\ \cos \lambda x \cos \mu x \sin \nu x &= \sin \{\lambda, \mu|\nu\}x = \sin \{\nu, \{\lambda, \mu|\}|\}x, \\ \sin \lambda x \sin \mu x \sin \nu x &= -\sin \{|\lambda, \mu, \nu\}x = -\sin \{\nu, \{|\lambda, \mu\}|\}x, \\ \sin \lambda x \sin \mu x \cos \nu x &= -\cos \{\nu|\lambda, \mu\}x = -\cos \{\{|\lambda, \mu\}, \nu|\}x. \end{aligned}$$

25. None of the twelve polynomials can be computed by linear recurrence, that is, by an ordinary difference equation of finite order, since otherwise an elliptic function would satisfy a linear differential equation of finite order and the first degree. We now give two symbolic linear recurrences by means of which, in conjunction with the foregoing addition theorems, the actual computation of all the polynomials can be effected systematically without reference to elliptic functions. Or, by using the results of §§ 30, 31, the polynomials can

be calculated directly from the recurrences. The addition theorems in any case shorten the labor but are not indispensable.

To find the recurrences for L, M of § 22 we proceed as if λ, μ were ordinaries, forming the equation in w whose roots are ϕ, ψ and multiplying the result by w^n, finally degrading exponents and replacing ordinary products by dot multiplications (§ 22). The formal algebraic details parallel those for the case of ordinaries, as given for example by Lucas 5, pp. 308–310, and may be omitted. For w = either L or M, and $n \geqq 0$, we find

$$w_{n+2} - 2\lambda_1 \cdot w_{n+1} + (\lambda_2 - \mu_2) \cdot w_n = 0,$$

with the initial values directly from the definitions of L, M,

$$L_0 = \lambda_0 \mu_0, \qquad L_1 = \lambda_1 \mu_0, \qquad M_0 = 0, \qquad M_1 = \lambda_0 \mu_1.$$

Thus for $n = 0$, $w = L$, we have

$$\begin{aligned} L_2 = 2\lambda_1 \cdot L_1 - (\lambda_2 - \mu_2) \cdot L_0 &= 2\lambda_1 \cdot \lambda_1 \mu_0 - (\lambda_2 - \mu_2) \cdot \lambda_0 \mu_0 \\ &= 2\lambda_2 \mu_0 - \lambda_2 \mu_0 + \lambda_0 \mu_2 = \lambda_2 \mu_0 + \lambda_0 \mu_2, \end{aligned}$$

which is correct. Put $n = 1$:

$$L_3 = 2\lambda_1 \cdot (\lambda_2 \mu_0 + \lambda_0 \mu_2) - (\lambda_2 - \mu_2) \cdot \lambda_1 \mu_0 = \lambda_3 \mu_0 + 3\lambda_1 \mu_2,$$

and so on.

Similarly if w is any one of A, B, C, D,

$$\begin{aligned} w_{n+4} - 4\lambda_1 \cdot w_{n+3} + 2(3\lambda_2 - \mu_2 - \nu_2) \cdot w_{n+2} - 4(\lambda_3 - \lambda_1 \mu_2 - \mu_2 \nu_1) \cdot w_{n+1} \\ + (\lambda_4 + \mu_4 + \nu_4 - 2\lambda_2 \mu_2 - 2\mu_2 \nu_2 - 2\nu_2 \lambda_2) \cdot w_n = 0; \end{aligned}$$

$$\begin{array}{ll} A_0 = \lambda_0 \mu_0 \nu_0, & B_0 = 0, \\ A_1 = \lambda_1 \mu_0 \nu_0, & B_1 = \lambda_0 \mu_0 \nu_1, \\ A_2 = \lambda_2 \mu_0 \nu_0 + \lambda_0 \mu_2 \nu_0 + \lambda_0 \mu_0 \nu_2, & B_2 = 2\lambda_1 \mu_0 \nu_1, \\ A_3 = \lambda_3 \mu_0 \nu_0 + 3\lambda_1 \mu_2 \nu_0 + 3\lambda_1 \mu_0 \nu_2; & B_3 = \lambda_0 \mu_0 \nu_3 + 3\lambda_2 \mu_0 \nu_1 + 3\lambda_0 \mu_2 \nu_1; \\ C_0 = 0, & D_0 = 0, \\ C_1 = 0, & D_1 = \lambda_0 \mu_1 \nu_0, \\ C_2 = 2\lambda_0 \mu_1 \nu_1, & D_2 = 2\lambda_1 \mu_1 \nu_0, \\ C_3 = 6\lambda_1 \mu_1 \nu_1; & D_3 = \lambda_0 \mu_3 \nu_0 + 3\lambda_2 \mu_1 \nu_0 + 3\lambda_0 \mu_1 \nu_2. \end{array}$$

Provided multiplications be interpreted as above, viz., as dot multiplications, the theory of such symbolic recurrences is obviously identical in its formal aspects with that of ordinary recurrences. Hence for L, M we have at once isomorphs of all the algebraic formulas developed by Lucas for his U, V. Thus, for example, L, M are expressible as symbolic continuants, cf. Lucas 4,

p. 193. Similarly A, B, C, D have symbolic properties completely analogous to the algebraic relations between any system of four independent solutions of

$$y_{n+4} + py_{n+3} + qy_{n+2} + ry_{n+1} + sy_n = 0;$$

in particular there are symbolic equivalents of the generalized continued fractions of Jacobi, Fürstenau, and others. This can be continued in the same way for functions of $4, 5, \cdots, n$ umbræ.

There is a more general aspect of the formulas of this part which is useful elsewhere. The development of any function of x which is even, odd, or arbitrary in x can be written in the symbolic forms

$$\cos \eta x, \quad \sin \zeta x, \quad \cos \eta x + i \sin \zeta x$$

respectively, and the coefficient of x^n in the product of any number of such functions can be most readily investigated by the symbolic trigonometry which we have sketched. When n functions are concerned the symbolic recurrences are of order n.

One special case of interest may be mentioned. As remarked in § 19, $(\Lambda|)^n$, $(|\Lambda')^n$, when exponents are interpreted as algebraic and the umbræ as ordinaries, are the symmetric functions of Kronecker 13, p. 385, which he took as his point of departure for deducing certain properties of the Bernoulli numbers. The step-by-step symbolic interpretation of Kronecker's analysis may be made without difficulty, and all his formulas translated into terms of umbræ instead of ordinaries. The formulas thus derived contain his as limiting cases. Bernoulli numbers enter Kronecker's formulas as coefficients in the expansion of hyperbolic tangents, and hence the manner in which they appear is quite distinct from that of the present discussion.

26. Corresponding to any product of elliptic functions there is a unique product of symbolic sines and cosines, and hence a unique $(\Lambda|\Lambda')^n$. From any elliptic identity we write down an identity between bar functions of general odd or even weight. When the identity contains ns x, cs x or ds x it is first multiplied throughout by the lowest power of x such that these functions can be replaced by x ns x, x cs x or x ds x respectively wherever they occur. Now the umbræ $\lambda, \mu, \cdots, \tau'$ in $(\Lambda|\Lambda')^n$ are its umbral factors (§ 19), and hence the isomorphism between elliptic and polynomial identities is complete.

III. Algebraic relations and congruences

27. For $z = 1$ the polynomials degenerate to Bernoulli and allied numbers, § 14. Let ω denote a complex cube root of unity. The cases $z = -\omega$, $z = \frac{1}{2}$, which by an obvious analogy may be called the equianharmonic and harmonic, are noted here in passing. They are characterized by the vanishing

of certain invariants (§ 4). Write

$$A_n(z),\ \alpha A_n(z),\ \beta A_n(z) \equiv A_n(z),\ A'_n(z),\ A''_n(z),$$

$$I_n(z) = \begin{vmatrix} A_n(z) & A'_n(z) & A''_n(z) \\ A'_n(z) & A''_n(z) & A_n(z) \\ A''_n(z) & A_n(z) & A'_n(z) \end{vmatrix}.$$

Then from the table in § 2,

$$I_n(z) = \begin{vmatrix} A_n(z) & A'_n(z) & A''_n(z) \\ \alpha A_n(z) & \alpha A'_n(z) & \alpha A''_n(z) \\ \beta A_n(z) & \beta A'_n(z) & \beta A''_n(z) \end{vmatrix},$$

and from the values of α', β' in § 1, it is evident that this vanishes when $z = \alpha' z$, $z = \beta' z$, $\alpha' z = \beta' z$, that is, when $z = -\omega$, $-\omega^2$. Hence

$$I_n(-\omega) = I_n(-\omega^2) = 0,$$

and therefore from the tables in §§ 13, 16 we have, when $u = -\omega$ or $-\omega^2$,

$$S^3_{2n+1}(u) + \alpha S^3_{2n+1}(u) + \beta S^3_{2n+1}(u) - 3S_{2n+1}(u)\alpha S_{2n+1}(u)\beta S_{2n+1}(u) = 0,$$

$$P^3_{2n+1}(u) + \alpha P^3_{2n+1}(u) + \beta P^3_{2n+1}(u) - 3P_{2n+1}(u)\alpha P_{2n+1}(u)\beta P_{2n+1}(u) = 0,$$

$$C^3_{2n}(u) + \alpha C^3_{2n}(u) + \beta C^3_{2n}(u) - 3C_{2n}(u)\alpha C_{2n}(u)\beta C_{2n}(u) = 0,$$

$$\gamma C^3_{2n}(u) + \epsilon C^3_{2n}(u) + \delta C^3_{2n}(u) - 3\gamma C_{2n}(u)\epsilon C_{2n}(u)\delta C_{2n}(u) = 0,$$

the fourth of which follows from the third on transforming by γ, or it is independently obvious from the last table in § 16. The product of the last two is another invariant of the same form which vanishes for the same values of u. This invariant corresponds to the entire last table in § 16, the determinant of which is a circulant of the sixth order, and hence, by a well known theorem of Glaisher, is expressible as a circulant of the third order. From the values of α'', β'' in § 1 and the form of $I_n(z)$ each of these invariants vanishes also when $z = 0,\ 1$. The harmonic invariants are written down in the same way.

28. Applying the first formulas of § 24 to the final set in § 17 we have, for $n > 0$,

$$\{C, C|\}^{2n} - \{|S, S\}^{2n} = 0,$$

$$\{\alpha C, \alpha C|\}^{2} + (1 - z)\{|\alpha S, \alpha S\}^{2n} = 0,$$

$$\{\beta C, \beta C|\}^{2n} + z\{|\beta S, \beta S\}^{2n} = 0,$$

$$\{\gamma C, \gamma C|\}^{2n} - z\{|S, S\}^{2n} = 0,$$

$$\{\delta C, \delta C|\}^{2n} + \{|\alpha S, \alpha S\}^{2n} = 0,$$

$$\{\epsilon C, \epsilon C|\}^{2n} - (1 - z)\{|\beta S, \beta S\}^{2n} = 0,$$

in which $C = C(z)$, and likewise for the rest. For $z = 1$ we get from § 14 the corresponding degenerate form, $(i = \sqrt{-1})$,

$$\{iE, iE|\}^{2n} - \{|H, H\}^{2n} = 0, \quad \text{or} \quad (-1)^n\{E, E|\}^{2n} = \{|H, H\}^{2n},$$

from the first or fourth of the above. The remaining degenerate forms are identities between powers of i. In full the degenerate relation is

$$(-1)^n \sum_{r=0}^{n} \binom{2n}{2r} E_{2n-2r} E_{2r} = \sum_{r=1}^{n} \binom{2n}{2r-1} H_{2n-2r+1} H_{2r-1}.$$

Similarly for the polynomials P with $n > 1$,

$$\{P, P|\}^{2n} - \{\beta P, \beta P|\}^{2n} = 0, \qquad \{P, P|\}^{2n} - \{\alpha P, \alpha P|\}^{2n} = 0,$$
$$2^{2n-2}\{B, B|\}^{2n} - \{R, R|\}^{2n} = 0,$$

the first of which corresponds to

$$x^2 \operatorname{ns}^2 x - x^2 \operatorname{ds}^2 x = x^2 k^2,$$

and the second comes from this as shown in § 17 by transforming with respect to α. The degenerate relation reduces by § 14 to

$$\sum_{r=0}^{n} \binom{2n}{2r} (2^{2n-2r-1} + 2^{2r-1} - 1) B_{2n-2r} B_{2r} = 0.$$

29. It is well known that all the algebra including the addition theorems of the elliptic functions follows from the square relations $\operatorname{cn}^2 x + \operatorname{sn}^2 x = 1$, etc., and the expressions for the derivatives of $\operatorname{sn} x$, $\operatorname{cn} x$. Hence all relations between the polynomials are implicit in § 28 and the formulas next given, which are written down from the expressions for the derivatives of $\operatorname{sn} x$, $\operatorname{cn} x$, $x \operatorname{ns} x$ and their transforms by § 16, or directly from the first member of each of the three sets as suggested at the end of § 17. The first formulas of § 24 are used for all, and the results may be checked at a glance by comparing with Glaisher 6, p. 92. Thus, the dot being as in § 22, we obtain from

$$\operatorname{sn}(x, k) = \sin S(z)x,$$

by differentiation,

$$\operatorname{cn}(x, k) \operatorname{dn}(x, k) = S_1(z) \cdot \cos S(z)x,$$

or

$$\cos C(z)x \cos \gamma C(z)x = S_1(z) \cdot \cos S(z)x,$$

which gives at once the first of the following:

$$S_{2n+1} = \{C, \gamma C|\}^{2n}, \qquad (2n+1) P_{2n+2} = -\{\alpha P, \beta P|\}^{2n+2},$$
$$\alpha S_{2n+1} = \{\alpha C, \delta C|\}^{2n}, \qquad (2n+1) \alpha P_{2n+2} = -\{P, \beta P|\}^{2n+2},$$
$$\beta S_{2n+1} = \{\beta C, \epsilon C|\}^{2n}; \qquad (2n+1) \beta P_{2n+2} = -\{P, \alpha P|\}^{2n+2};$$
$$C_{2n+2} = \{S, \gamma C|\}^{2n+1}, \qquad \gamma C_{2n+2} = z\{S, C|\}^{2n+1},$$
$$\alpha C_{2n+2} = -(1 - z)\{\alpha S, \delta C|\}^{2n+1}, \qquad \delta C_{2n+2} = -\{\alpha S, \alpha C|\}^{2n+1},$$
$$\beta C_{2n+2} = -z\{\beta S, \epsilon C|\}^{2n+1}; \qquad \epsilon C_{2n+2} = (1 - z)\{\beta S, \beta C|\}^{2n+1},$$

in all of which $n \geqq 0$. By means of the L, M addition theorems (§ 22) and recurrences (§ 25) all of the polynomials can be calculated successively, and without excessive labor, from the initial values for $n = 0, 1$ (which are given by the definitions). Putting $z = 1$ as before we find the degenerate forms:

$$H_{2n+1} = (-1)^n\{E, E|\}^{2n}, \qquad (2n+1)2^{2n}B_{2n+2} = -\{R, R|\}^{2n+2},$$

$$E_{2n+2} = (-1)^{n+1}\{H, iE|\}^{2n}, \qquad (2n+1)R_{2n+2} = -\{2B, R|\}^{2n+2};$$

whence, by comparison with those in § 29,

$$H_{2n+1} = \{|H, H\}^{2n}, \qquad (2n+1)B_{2n+2} = -\{B, B|\}^{2n+2}.$$

30. As a last example of relations involving bar functions of not more than two umbræ we take the identities which express the elliptic functions as products of two others, and those between the functions and their reciprocals. The argument is z, as before.

$$\begin{aligned} S_{2n+1} &= \{\alpha S, C|\}^{2n+1} = \{\beta S, \gamma C|\}^{2n+1}, \\ \alpha S_{2n+1} &= \{\beta S, \alpha C|\}^{2n+1} = \{S, \delta C|\}^{2n+1}, \\ \beta S_{2n+1} &= \{S, \beta C|\}^{2n+1} = \{\alpha S, \epsilon C|\}^{2n+1}; \\ P_{2n} &= \{\alpha P, \delta C|\}^{2n} = \{\beta P, \beta C|\}^{2n}, \\ \alpha P_{2n} &= \{\beta P, \epsilon C|\}^{2n} = \{P, C|\}^{2n}, \\ \beta P_{2n} &= \{P, \gamma C|\}^{2n} = \{\alpha P, \alpha C|\}^{2n}; \end{aligned}$$

$$\begin{aligned} (2n+1)C_{2n} &= \{S, \alpha P|\}^{2n+1}, & C_{2n} &= \{\gamma C, \epsilon C|\}^{2n}, \\ (2n+1)\alpha C_{2n} &= \{\alpha S, \beta P|\}^{2n+1}, & \alpha C_{2n} &= \{\delta C, \gamma C|\}^{2n}, \\ (2n+1)\beta C_{2n} &= \{\beta S, P|\}^{2n+1}, & \beta C_{2n} &= \{\epsilon C, \delta C|\}^{2n}, \\ (2n+1)\gamma C_{2n} &= \{S, \beta P|\}^{2n+1}, & \gamma C_{2n} &= \{C, \alpha C|\}^{2n}, \\ (2n+1)\delta C_{2n} &= \{\alpha S, P|\}^{2n+1}, & \delta C_{2n} &= \{\alpha C, \beta C|\}^{2n}, \\ (2n+1)\epsilon C_{2n} &= \{\beta S, \alpha P|\}^{2n+1}, & \epsilon C_{2n} &= \{\beta C, C|\}^{2n}; \end{aligned}$$

whence the degenerate forms,

$$\begin{aligned} H_{2n+1} &= (-1)^n\{1, E|\}^{2n+1}, & \{1, E|\}^{2n} &= 0, \quad n > 0, \\ 1 &= (-1)^n\{H, i|\}^{2n+1}, & (2n+1)E_{2n} &= 2(-1)^n\{H, iR|\}^{2n+1}, \\ 2^{2n-1}B_{2n} &= \{R, 1|\}^{2n}, & \{1, R|\}^{2n+1} &= 0, \quad n > 0, \\ 2n+1 &= \{1, 2B|\}^{2n+1}, & 2R_{2n} &= \{2B, E|\}^{2n}. \end{aligned}$$

31. The simplest general relations between polynomials of one kind and of several ranks are found from the differential equations for the elliptic functions. Let λ, μ, ν denote umbræ and consider functions v, t, w, u of x such that

$$\begin{aligned} v &= \cos \lambda x, & v'' &= av + 2bv^3, \\ t &= \sin \mu x, & t'' &= ct + 2dt^3, \\ w &= xu = \cos \nu x, & u'' &= gu + 2hu^3, \end{aligned}$$

in which a, b, c, d, g, h are independent of x, and double accents denote second derivatives with respect to x. Then v, t, u are elliptic functions, and hence λ, μ, ν polynomials in z (= the square of the modulus). From the differential equations for v, t, w we get by substituting the symbolic trigonometric equivalents of v, t, w, performing the differentiations and equating coefficients of like powers of x,

$$\lambda_{2n+2} + a\lambda_{2n} + 2b\{\lambda, \lambda, \lambda|\}^{2n} = 0,$$

$$\mu_{2n+3} + c\mu_{2n+1} - 2d\{|\mu, \mu, \mu\}^{2n+1} = 0,$$

$$n(2n+1)\nu_{2n+2} + (n+1)(2n+1)g\nu_{2n} - h\{\nu, \nu, \nu|\}^{2n+2} = 0,$$

in which we have used the last formulas of § 24 for the coefficients* in v^3, t^3, w^3. Knowing the values of the constants in the cases $\lambda = C$, $\mu = S$, $\nu = P$ from Glaisher 6, p. 122, we write down the rest of the following table by the methods of § 17. The argument is z throughout.

$$\begin{array}{lll} (\lambda, a, b) = & (\mu, c, d) = & (\nu, g, h) = \\ (C, 2z-1, -z), & (S, -1-z, z), & (P, -1-z, 1), \\ (\alpha C, -1-z, 1), & (\alpha S, 2-z, 1-z), & (\alpha P, 2-z, 1), \\ (C, 2-z, -1+z), & (\beta S, 2z-1, -z+z^2); & (\beta P, 2z-1, 1). \\ (\gamma C, 2-z, -1), & & \\ (\delta C, 2z-1, 1-z), & & \\ (\epsilon C, -1-z, z); & & \end{array}$$

We omit the degenerate forms. The bar functions of λ, μ, ν may be computed by the addition theorems of § 23 and the recurrences of § 25, and hence each of the polynomials for all ranks can be calculated independently of the remaining polynomials. In using §§ 23, 25 we perform all operations for three distinct umbræ, putting these equal to one another in the successive final steps (see § 18). We must pass on to a brief discussion of the congruences for prime moduli.

---

* The differential equation for w is $x^2 w'' = 2xw' + (gx^2 - 2)w + hu^3$.

32. Henceforth p denotes an odd prime > 0. (Some of the congruences are also valid modulo 2, but this case is of such slight interest that we ignore it.) The theorem of arithmetic which gives the congruences modulo p is due to Lucas 4, pp. 229–230:

$$\binom{m}{n} \equiv \binom{m_1}{n_1}\binom{m_1'}{n_1'} \qquad \mod p,$$

$$m = m_1 p + m_1', \qquad n = n_1 p + n_1', \qquad 0 \leqq m_1', \quad n_1' < p.$$

We shall in future omit the "mod p" in writing congruences, and use "$\equiv$" only in the sense of congruence modulo p. From Lucas' theorem it is easy to infer that, λ, μ denoting umbræ of integers and the significance of the dot being as in § 22,

$$(\lambda \pm \mu)^{rp+s} \equiv (\lambda^p \pm \mu^p)^r \cdot (\lambda \pm \mu)^s, \qquad 0 \leqq s < p,$$

the upper signs or the lower being taken throughout. By repeated application of this we find the following. Let

$$N = r_n p^n + r_{n-1} p^{n-1} + r_{n-2} p^{n-2} + \cdots + r_1 p_1 + r_0$$

be the (unique) expression of N in the scale of p, so that $0 \leqq r_j < p$ $(j = 0, 1, \cdots, n)$. Then

$$(\lambda + \mu)^N \equiv (\lambda^{p^n} + \mu^{p^n})^{r_n} \cdot (\lambda^{p^{n-1}} + \mu^{p^{n-1}})^{r_{n-1}} \cdots (\lambda^p + \mu^p)^{r_1} \cdot (\lambda + \mu)^{r_0},$$

$$(\lambda - \mu)^N \equiv (\lambda^{p^n} - \mu^{p^n})^{r_n} \cdot (\lambda^{p^{n-1}} - \mu^{p^{n-1}})^{r_{n-1}} \cdots (\lambda^p - \mu^p)^{r_1} \cdot (\lambda - \mu)^{r_0}.$$

Denote the right-hand members of these congruences by

$$\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} \pm \mu^{p^{n-j}})^{r_{n-j}}$$

respectively, in which the accented Π' indicates that the products of the several factors are to be performed as dot multiplications. Then

$$\{\lambda, \mu|\}^N \equiv \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} + \mu^{p^{n-j}})^{r_{n-j}} + \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} - \mu^{p^{n-j}})^{r_{n-j}},$$

$$\{|\lambda, \mu\}^N \equiv \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} + \mu^{p^{n-j}})^{r_{n-j}} - \tfrac{1}{2}\prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} - \mu^{p^{n-j}})^{r_{n-j}}.$$

The distributed *non-symbolic* equivalent of each right-hand member after the performance of all dot multiplications and the degradation of all exponents is a quadratic form with positive integral coefficients of the form

$$\binom{r_n}{r_n'}\binom{r_{n-1}}{r_{n-1}'}\cdots\binom{r_0}{r_0'} \qquad (r_n' \leqq r_n, \quad r_{n-1}' \leqq r_{n-1}, \quad \cdots, \quad r_0' \leqq r_0),$$

and the term of which this is the coefficient is

$$\lambda_{N-N'}\,\mu_{N'} \qquad (N' = r'_n p^n + r'_{n-1}p^{n-1} + \cdots + r'_0).$$

That the coefficients are positive is evident from the left-hand members in which all coefficients are such; the coefficients on the right are the positive residues modulo p of those on the left.

If $\rho_N = \{\lambda, \mu|\}^N$, or if $\rho_N = \{|\lambda, \mu\}^N$, the above congruences are applicable in an obvious manner, equalities of this sort having been considered in the preceding sections. From § 15 it is clear that any congruence between polynomials (the P set does not immediately enter the discussion since the coefficients are not integers) is equivalent to a system of congruences between the linear functions of the form in § 3 into which H, E and their anharmonic transforms (the degenerate forms of the corresponding polynomials) for different ranks are partitioned.

Two important special cases of the general formula in this section are

$$\{\lambda, \mu|\}^p \equiv \lambda_p\,\mu_0, \qquad \{|\lambda, \mu\}^p \equiv \lambda_0\,\mu_p.$$

Hence in all cases we have the residue of a bar function of two integral umbræ when the rank of the function is equal to or greater than the modulus. The case of two umbræ will be completed for a prime modulus when we find, as next, the residue when the rank is less than the modulus.

33. Restating another theorem of Lucas 4, p. 229, in a form adapted to our purpose we have

$$\binom{p-1}{n} \equiv (-1)^n, \qquad \binom{p-r}{n} \equiv (-1)^n\binom{n+r-1}{r-1}$$

$$(0 < r < p, \quad n < p - r).$$

Hence for $r > 0$,

$$(\lambda + \mu)^{p-r} \equiv \sum_{s=0}^{p-r}(-1)^s\binom{r+s-1}{s}\lambda_{p-r-s}\,\mu_s,$$

$$(\lambda - \mu)^{p-r} \equiv \sum_{s=0}^{p-r}\binom{r+s-1}{s}\lambda_{p-r-s}\,\mu_s;$$

and hence, including all cases (cf. § 24), for $h \geqq 0$,

$$\{\lambda, \mu|\}^{p-2h-1} \equiv \sum_{s=0}^{(p-2h-1)/2}\binom{2h+2s}{2s}\lambda_{p-2h-2s-1}\,\mu_{2s},$$

$$\{\lambda, \mu|\}^{p-2h-2} \equiv \sum_{s=0}^{(p-2h-3)/2}\binom{2h+2s+1}{2s}\lambda_{p-2h-2s-2}\,\mu_{2s},$$

$$\{|\lambda, \mu\}^{p-2h-1} \equiv - \sum_{s=0}^{(p-2h-1)/2} \binom{2h+2s-1}{2s-1} \lambda_{p-2h-2s}\, \mu_{2s-1}.$$

34. The congruences in § 33 admit of immediate generalization to any number m of umbræ $\lambda, \rho, \cdots, \mu$. It is easily seen in the same way, or as a consequence of the fundamental formula of § 33,

$$(\lambda + \mu)^{rp+s} \equiv (\lambda^p + \mu^p)^r \cdot (\lambda + \mu)^s,$$

that we have

$$(\lambda + \rho + \cdots + \mu)^{rp+s} \equiv (\lambda^p + \rho^p + \cdots + \mu^p)^r \cdot (\lambda + \rho + \cdots + \mu)^s;$$

and hence in the previous notation,

$$(\lambda + \rho + \cdots + \mu)^N \equiv \prod_{j=0}^{n}{}' (\lambda^{p^{n-j}} + \rho^{p^{n-j}} + \cdots + \mu^{p^{n-j}})^{r_{n-j}},$$

with $2^m - 1$ similar congruences in which the signs of some or all of $\lambda, \rho, \cdots, \mu$ are changed throughout. From these the congruences for the general bar functions of § 19 follow at once. The particular case $m = 3$ is of interest in connection with the formulas of § 32, but the length of this paper precludes further discussion. From the general formulas given there is no difficulty in writing down the congruences for bar functions of two or three letters from the algebraic relations developed earlier in the paper. For $z = 1$ some of the degenerate cases are well known, and hence provide checks. The other degenerate cases may be checked directly in a similar manner.

References

1. J. Blissard: *Theory of generic equations*, Quarterly Journal of Mathematics, vol. 4 (1861), pp. 279–305; vol. 5 (1862), pp. 58–75, 185–208.
2. J. Blissard: *Examples of the use and application of the representative notation*, ibid., vol. 6 (1864), pp. 49–64.
3. J. Blissard: *Researches in analysis*, ibid., vol. 7 (1865), pp. 155–170, 223–226.
4. E. Lucas: *Théorie des fonctions numériques simplement périodiques*, American Journal of Mathematics, vol. 1 (1878), pp. 184–238, 289–321.
5. E. Lucas: *Théorie des Nombres*, 1891.
6. J. W. L. Glaisher: *On elliptic functions*, Messenger of Mathematics, vol. 11 (1881–1882), pp. 81–95, 120–138.
7. J. W. L. Glaisher: *On the Bernoullian function*, Quarterly Journal of Mathematics, vol. 29 (1898), pp. 1-168 (especially pp. 122–168).
8. J. W. L. Glaisher: *On a class of relations connecting any n consecutive Bernoullian functions*, ibid., vol. 42 (1911), pp. 86–157.
9. A. Cayley: *An Elementary Treatise on Elliptic Functions*, Edition 2 (1895).
10. C. Hermite: *Remarques sur le développement de* cos am x, Paris Comptes Rendus, vol. 67 (1863), p. 613; Journal de Mathématiques pures et appliquées, (2) vol. 9 (1864), p. 289; Œuvres, vol. 2, pp. 265–270.
11. C. Hermite: *Sur les développements de* $F(x) = \text{sn}^a\, x\, \text{cn}^b\, x\, \text{dn}^c\, x$ *où les exposants sont entiers*, Académie royale des Sciences de Stockholm, Bihang III, No. 10 (1875), pp. 3–10; Œuvres, vol. 3, pp. 222–235.

12. C. HERMITE: *Sur le développement des fonctions elliptiques*, Journal für die reine und angewandte Mathematik, vol. 81 (1876), pp. 220–228; Œuvres, vol. 3, pp. 236–245.

13. L. KRONECKER: *Sur quelques fonctions symétriques et sur les nombres de Bernoulli*, Journal de Mathématiques pures et appliquées, (2) vol. 1 (1856), pp. 385–391.

14. L. KRONECKER: *Ueber die Bernoullischen Zahlen*, Journal für die reine und angewandte Mathematik, vol. 94 (1883), pp. 268–269.

UNIVERSITY OF WASHINGTON,
SEATTLE, WASH.

NEW PROPERTIES OF ALL REAL FUNCTIONS*

BY

HENRY BLUMBERG

Introduction

In a former paper† the author communicated a number of properties of every real function $f(x)$; these were stated in terms of the successive saltus functions associated with a given function. The present paper makes no use of the saltus functions, and the new properties are direct qualifications of $f(x)$. Since $f(x)$ is entirely unrestricted,—except, of course, that it is defined ‡ and therefore finite for every real x—these qualifications are consequences of nothing else than that $f(x)$ is a function. A new light is thus thrown upon the nature of a function.

The new properties are of two types, descriptive and metric; the former are concerned with density, and the latter with measure (Lebesgue).

For the sake of greater concreteness of exposition, we shall discuss, for the most part, planar sets and real functions of two real variables.

1. Descriptive properties

We shall say that a planar set S is an *I-region* (= open set) if every point of S is an inner point of S; i.e., no point of S is the limit of a sequence of points not in S.

We shall deal with binary relations $\mathfrak{R}$ between I-regions and points. $I\mathfrak{R}P$ shall mean that the I-region I has the relation $\mathfrak{R}$ to the point P. The relation $\mathfrak{R}$ is said to be *closed*, if the relationships $I\mathfrak{R}P_n$ and $\lim_{n\to\infty} P_n = P$ imply $I\mathfrak{R}P$. By a *neighborhood* of a point P, we understand an I-region containing P; by a *partial neighborhood* of P, an I-region of which P is an inner or boundary point. A neighborhood of P is, therefore, also a partial neighborhood of P. We have the following

LEMMA I. *If $\mathfrak{R}$ is a closed relation, then the points for which (a) $N\mathfrak{R}P$ for every neighborhood N of P, and (b) a partial neighborhood $N_<$ exists such that $\overline{N_< \mathfrak{R}P}$ (i.e., $N_< \mathfrak{R}P$ is false) constitute a non-dense (i.e., nowhere dense) set M.*

* Presented to the Society March 29, 1919, and December 30, 1920.

† *Certain general properties of functions*, Annals of Mathematics, vol. 18 (1917), p. 147.

‡ Even this restriction may be partially dispensed with; cf. Section 2.

Proof: Since $\mathfrak{R}$ is closed and $\overline{N_< \mathfrak{R} P}$, it follows that no sequence of points P_n exists such that $\lim P_n = P$ and $\overline{N_< \mathfrak{R} P_n}$ for every n. Hence a neighborhood N of P exists such that $\overline{N_< \mathfrak{R} A}$ for every point A of N. In particular $\overline{N_< \mathfrak{R} B}$ for every point B of $NN_<$, the set of points common to N and $N_<$. Since $N_<$ is a neighborhood of B, it follows that no B belongs to M. Every neighborhood of a point P of M thus contains a subset, which is an I-region, every point of which is outside of M.

Now let $z = f(x, y)$ be any given real function of the two real variables x and y, defined for every point of the plane Π, which we take to be the XY plane of a cartesian system of coördinates. We shall write also $z = f(P)$, where $P = (x, y)$ ranges over Π. For every pair of real numbers (x, y), there is a single number $f(x, y)$; otherwise f is unrestricted. We define as follows the relation $\mathfrak{R}_{r_1 r_2}$, where r_1 and r_2 are two real numbers and $r_1 < r_2$: *If P is a point and I an I-region of Π, then $I \mathfrak{R}_{r_1 r_2} P$, if and only if an infinite sequence of points P_n of I exists, such that*

$$\lim P_n = P, \qquad \lim f(P_n) \textit{ exists,}$$

and

$$r_1 \leqq \lim f(P_n) \leqq r_2.$$

It follows that $\mathfrak{R}_{r_1 r_2}$ is closed. For suppose that $\lim P^{(n)} = P$ and $I \mathfrak{R}_{r_1 r_2} P^{(n)}$ for every n. Then, for every n, there exists a sequence

$$\{P_m^{(n)}\}, \qquad m = 1, 2, \cdots,$$

of points of I such that $\lim_{m \to \infty} P_m^{(n)} = P^{(n)}$ and $r_1 \leqq \lim_{m \to \infty} f(P_m^{(n)}) \leqq r_2$. From the $P_m^{(n)}$, we may, in view of the preceding relations, select a sequence of points $\{Q_n\}$, $n = 1, 2, \cdots$, such that

$$\lim Q_n = P \qquad \text{and} \qquad r_1 \leqq \liminf f(Q_n) \leqq \limsup f(Q_n) \leqq r_2;$$

and from the sequence $\{Q_n\}$ a subsequence $\{P_n\}$ such that $\lim f(P_n)$ exists. We then have

$$\lim P_n = P \qquad \text{and} \qquad r_1 \leqq \lim f(P_n) \leqq r_2.$$

$\mathfrak{R}_{r_1 r_2}$ is therefore closed.

By the use of Lemma I, we may thus conclude that for every function $f(x, y)$ and every pair of numbers $r_1 < r_2$, the points P of the XY plane for which (*a*) $N \mathfrak{R}_{r_1 r_2} P$ for every neighborhood N of P and (*b*) $\overline{N_< \mathfrak{R}_{r_1 r_2} P}$ for some partial neighborhood $N_<$ of P constitute a non-dense set $T_{r_1 r_2}$. Let T be the sum of all the sets $T_{r_1 r_2}$, r_1 and r_2 $(r_1 < r_2)$ ranging independently over all the rational numbers. T, being the sum of a denumerable number of non-dense sets, is exhaustible* (i.e., of first category according to Baire).

*For the terminology cf. Denjoy, Journal de Mathématiques, ser. 7, vol. 1 (1915), pp. 122–125.

We shall now use this property of T to obtain a property of all functions. For this purpose, we introduce the notion of *dense approach.* The function $f(x, y)$ is said to be *densely approached at the point* (ξ, η), or in other words, the point (ξ, η, ζ), $\zeta = f(\xi, \eta)$, of the "surface" $z = f(x, y)$ is said to be densely approached, if for every positive ϵ there exists a planar neighborhood N of (ξ, η), such that the points (x, y) of N for which $|f(x, y) - f(\xi, \eta)| < \epsilon$ form a dense set in N.

We have the following

THEOREM I. *For every real function $f(x, y)$ whatsoever, the points of the surface $z = f(x, y)$ that are densely approached form a residual set (= complement of an exhaustible set). Conversely, given any residual set R whatsoever, a function $f(x, y)$ exists that is densely approached at and only at the points of R.*

*Proof.** The points of Π'—i.e., of the surface $z = f(x, y)$—are either isolated or not. The isolated points of Π' form a denumerable set; therefore the points of Π whose corresponding points of Π' are isolated form a denumerable and therefore an exhaustible set. It is consequently sufficient to prove that the points P of Π, for which P' is a limit point of Π' but is not densely approached by Π', form an exhaustible set. Since P' is not densely approached, there exists an ϵ such that for every neighborhood N of P, the points Q of N for which $|f(Q) - f(P)| < \epsilon$ are not dense in N; i.e., an I-region exists in every neighborhood N of P such that $|f(Q) - f(P)| \geqq \epsilon$ for every point Q of the I-region. Hence there exists a partial neighborhood $N_<$ of P such that the inequality $|f(P) - f(Q)| \geqq \epsilon$ holds for every point Q of $N_<$. Let now r_1 be a rational number between $f(P) - \epsilon$ and $f(P)$; and r_2 a rational number between $f(P)$ and $f(P) + \epsilon$. Then it follows from the inequality $|f(Q) - f(P)| \geqq \epsilon$ for all points Q of $N_<$ that the set $\{Q'\}_{Q \text{ in } N_<}$—i.e., the set of points of Π' corresponding to the points Q of $N_<$—has no point (ξ, η, z) as a limit with $r_1 \leqq z \leqq r_2$, where $(\xi, \eta) = P$. Hence $\overline{N_<} \, \mathfrak{R}_{r_1 r_2} \, P$. On the other hand, P' is a limit point of Π' and hence of every N' where N is a neighborhood of P; hence N' has (ξ, η, ζ), with $r_1 < \zeta = f(\xi, \eta) < r_2$, as a limit point, and P is therefore a point of the non-dense set $T_{r_1 r_2}$ associated with the closed relation $\mathfrak{R}_{r_1 r_2}$, and hence of the exhaustible set T which is the sum of all the $T_{r_1 r_2}$, r_1 and r_2 ranging independently over all the rational numbers. The points of Π' that are not densely approached thus constitute an exhaustible set.

The proof of the converse is immediate. For let E be the exhaustible set complementary to the given residual set R. We may write

$$E = E_1 + E_2 + \cdots E_n + \cdots$$

* If S is any subset of Π, we shall understand by S' the set of points of the surface $z = f(x, y)$ corresponding to the points of S. Thus Π' represents the totality of surface points. Analogously, P' will represent the surface point corresponding to P.

where the E_n's are non-dense in Π and no pair of E_n's have common points. Let

$$f(P) = \frac{1}{n}, \quad \text{if} \quad P \text{ is in } E_n,$$

and

$$f(P) = 2, \quad \text{if} \quad P \text{ is in } R.$$

It is then clear that $f(P)$ is densely approached, if and only if P is in R.

Theorem I shows what a remarkable degree of "regularity" every function possesses. This may be better realized, perhaps, by using the following equivalent definition of dense approach. The function $f(x, y)$ is said to be densely approached at P, if for every partial neighborhood $N_<$ of P, $N'_<$ has P' as a limit. We thus get the following equivalent theorem—we omit a restatement of the converse—which shows a kind of "microscopic symmetry" in the structure of the surface $z = f(x, y)$ for unconditioned f.

THEOREM I'. *With every function $f(x, y)$ whatsoever, there is associated a residual set R—dependent on f—of the XY plane such that if $P = (\xi, \eta)$ is a point of R, and $N_<$ a partial neighborhood of P, the set $N'_<$, which consists of the points of the surface $z = f(x, y)$ that correspond to the points of $N_<$, has $(\xi, \eta, f(\xi, \eta))$ as a limit point.*

We pass now to further discriminations in the manner of "approach." The function f is said to be *inexhaustibly approached* at the point P of Π, or in other words, P' is inexhaustibly approached by Π', if every neighborhood of P contains, for every $\epsilon > 0$, an inexhaustible set of points—i.e., a set that is not exhaustible,—at which f differs from $f(P)$ by less than ϵ.

The function f is said to be *exhaustibly approached*—we then say also that the point $P' = (x, y, f(x, y))$ is exhaustibly approached by Π'—at the point $P = (x, y)$, if it is not inexhaustibly approached at the point; in other words, if a neighborhood N of (x, y) and a number $\epsilon > 0$ exist, such that the points of N where f differs from $f(x, y)$ by less than ϵ form an exhaustible set. If M is any planar set, we shall use, in connection with approach, the expression "via M" to indicate that (x, y) is restricted to range in M. Thus "f is inexhaustibly approached at P via M" means that for every neighborhood N of P and every $\epsilon > 0$, the set MN, which is the aggregate of points common to M and to N, contains an inexhaustible set of points at which f differs from $f(P)$ by less than ϵ. The following definition for exhaustible approach is equivalent to the one above: The point P' of Π' is exhaustibly approached, if a sphere S exists with P' as center such that the points of Π corresponding to the points of $\Pi'S$ form an exhaustible set.

We have the following

THEOREM II. *For every function $f(x, y)$, there exists in the XY plane a*

*residual set R, dependent on f, such that if P is a point of R, and $N_<$ a partial neighborhood of P, the function f is inexhaustibly, and therefore densely, approached * at P via $RN_<$.*

Proof. We first note that *the points of* Π *where f is exhaustibly approached form an exhaustible set E_1.* For space contains a denumerable dense set

$$A = \{A_1, A_2, \cdots, A_m, \cdots\}.$$

Associate with each A_m a sequence of spheres S_{mn}, $n = 1, 2, \cdots$, with A_m as center and $1/n$ as radius. Blacken every sphere S_{mn}—the interior as well as the boundary—which is such that the totality of points of Π' contained in it corresponds to an exhaustible subset of Π. Let E_1 be the subset of Π constituted by the points which correspond to points of Π' that lie in one or more black spheres. Since the number of black spheres is at most denumerable and each black sphere contributes an exhaustible set to E_1, it follows that E_1 is exhaustible. If P is a point at which f is exhaustibly approached, P' is in at least one black sphere; hence P belongs to E_1.

Now suppose we remove from Π the set E_1 of points at which f is exhaustibly approached, thus obtaining the set $R_1 = \Pi - E_1$ of points where f is inexhaustibly approached. Since the values of f in an exhaustible set cannot affect the property of inexhaustible approach, every point of R_1 is inexhaustibly approached via R_1. Furthermore the points of R_1, at which f is densely approached via R_1, form a residual set of Π. For assume for the present that f is bounded,—we shall later drop this restriction. Let k be any number greater than the least upper bound of f. Let the function $g(x, y)$ be equal to k at the points of E_1 and equal to $f(x, y)$ at the points of R_1. Since the set of points at which $g(x, y)$ is densely approached forms a residual set R_g of which only an exhaustible subset can lie in E_1, there must be a residual set of Π in R_1 at the points of which g is densely approached. This dense approach is clearly valid via R_1 and for the function f. Let E_2 be the exhaustible subset of points of R_1 at which f is not densely approached via R_1. It will now be seen that at every point of the set $R = \Pi - E_1 - E_2$, f is densely and inexhaustibly approached via R, and furthermore that f is inexhaustibly approached at every point P of R via $N_< R$, the common part of $N_<$ and R, where $N_<$ is any partial neighborhood of P. Every point of R is inexhaustibly approached via R_1 and hence via R, since an exhaustible set has no effect upon inexhaustible approach. Moreover every point P of R is densely approached via R_1; i.e., if $N_<$ is a partial neighborhood of P, and ϵ a positive number, there exists a point Q of $N_< R_1$ at which f differs from $f(P)$ by less than ϵ. Since, however, f is inexhaustibly approached at Q via R, and since R

* f is densely approached at P via M, which implicitly has P as limit, if for every partial neighborhood $N_<$ of P such that $N_< M$ has P as limit, the set $(N_< M)'$, i.e., the set of surface points corresponding to the points of $N_< M$, has P' as a limit.

differs from R_1 by an exhaustible set, it follows that Q' is a limit point of R', and therefore $N_<$ contains a point of R at which f differs from $f(P)$ by less than ϵ. P is thus densely and inexhaustibly approached via $N_< R$.

If f is unbounded, we use the transformation

$$\bar{f}(x, y) = \frac{f(x, y)}{1 + |f(x, y)|}.$$

The new function $\bar{f}$ is bounded, but the properties of dense and of exhaustible and of inexhaustible approach are preserved by this transformation. We may therefore drop the restriction made above in reference to boundedness.

THEOREM III. *With every function $f(x, y)$ there is associated (not uniquely, however) a dense set D of the XY plane such that $f(x, y)$ is continuous, if (x, y) ranges over D.*

Proof. We shall say that a set of points is "*ϵ-spaced*" (in the plane Π) if every circle of radius ϵ contains at least one point of the set. Let $\epsilon_1, \epsilon_2, \epsilon_3, \cdots$ be a decreasing sequence of positive numbers with $\lim \epsilon_m = 0$. Let K_1 be any set that is everywhere dense in Π and at the points of which $f(x, y)$ is densely approached via K_1; such a set exists according to Theorem II. Since K_1 is everywhere dense in Π, we may select in K_1 a subset

$$D_1 = \{P_{11}, P_{12}, \cdots, P_{1n}, \cdots\},$$

which is isolated (i.e., no point of D is a limit point of D) and ϵ_1-spaced. We enclose each P_{1n} in a circle C_{1n} in such a way that (a) no two C_{1n}'s overlap—for this purpose it is sufficient to make the radius of C_{1n} less than one half the greatest lower bound of the distances from P_{1n} to the other points of D_1; and that (b) C_{1n} contains a dense subset of K_1 at every point of which the value of f differs from $f(P_{1n})$ by less than ϵ_1—the existence of such a subset for a sufficiently small radius of C_{1n} is guaranteed by the fact that f is densely approached at P_{1n} via K_1. Denote by J_{1n} the set of all the points in K_1 that are in the interior of C_{1n} and at which f differs from $f(P_{1n})$ by less than ϵ_1, and by K_2 the subset of K_1 constituted by all the points of all the J_{1n} and the points of K_1 that lie in the interior of $\Pi - \sum_{n=1}^{\infty} C_{1n}$. At every point Q of K_2 f is densely approached via K_2. For if Q is an interior point of $\Pi - \sum C_{1n}$, this property is obvious. If, however, Q is in C_{1n}, let $|f(Q) - f(P_{1n})| = \epsilon_1 - \delta$, $\delta > 0$. f is densely approached at Q via points of K_1 at which f differs from $f(Q)$ by less than δ; that is, via points where f differs from $f(P_{1n})$ by less than ϵ_1; hence via K_2. Since K_2 is everywhere dense and f is densely approached via K_2 at every one of its points, we may treat it as we did K_1. We select an isolated, ϵ_2-spaced subset $D_2 = \{P_{21}, P_{22}, \cdots, P_{2n}, \cdots\}$ of K_2 such that D_2 contains every point of D_1. We now enclose each P_{2n} in a circle C_{2n} in such a way that (a) C_{2n} contains a dense subset of points of K_2

where the value of f differs from $f(P_{2n})$ by less than ϵ_2; (*b*) no two C_{2n} overlap; and (*c*) C_{2n} lies either entirely within a C_{1n} or entirely outside of all C_{1n}. Denote by J_{2n} the set of all the points of $C_{2n}\, K_2$ at which f differs from $f(P_{2n})$ by less than ϵ_2, and by K_3 the subset of K_2 constituted by all the points of all the J_{2n} and the points of K_2, lying neither in the interior nor on the boundary of any C_{2n}. The set K_3, like K_2, is everywhere dense and such that at every point of it f is densely approached via K_3, and the process may therefore be continued. The set

$$D_m = \{P_{m1}, P_{m2}, \cdots, P_{mn}, \cdots\}$$

is an ϵ_m-spaced, isolated subset of K_m and a superset of D_{m-1}. The circle C_{mn} enclosing P_{mn} (*a*) contains a dense subset of points of K_m where f differs from $f(P_{mn})$ by less than ϵ_m, (*b*) has no points in common with C_{mp} for $p \neq n$ and (*c*) lies entirely outside of all C_{pq}, $p < m$, if P_{mn} is outside of all C_{pq} for $p < m$, and entirely inside every C_{pq}, $p < m$, containing P_{mn}.

We now define the set D of the theorem to be the sum of the D_m's ($m = 1, 2, \cdots$). That D is dense follows from the fact that D_m is ϵ_m-spaced. If P_{mn} is a point of D, it lies in every circle of a sequence of circles C_{mn}, $C_{m+1,\, n^{(1)}}, \cdots$, such that at the points of $C_{m+p,\, n^{(p)}}$ that belong to D the function f differs from $f(P_{mn})$ by less than ϵ_{m+p}. This proves the asserted continuity.

2. Generalizations

So far we have dealt only with real functions of two real variables. It is apparent that the considerations apply equally well to functions of a single variable and to functions of n real variables. The results are, however, essentially of a still more general character. Without entering upon extreme refinements of generalization we may note that the arguments in the preceding section are substantially valid for every function $f(P)$, defined in a set $\mathfrak{S}$ satisfying the following conditions:

(1) $\mathfrak{S}$ is a metric;* that is to say, with every pair of elements P and Q of $\mathfrak{S}$ there is associated a non-negative, real number $\overline{PQ}$ (Fréchet's *écart*) in such a way that if P, Q, and R are any three elements of $\mathfrak{S}$, then

(*a*) $\overline{PQ} = \overline{QP}$;

(*b*) $\overline{PQ} = 0$, when and only when $P = Q$; and

(*c*) $\overline{PQ} + \overline{QR} \geqq \overline{PQ}$.

(2) $\mathfrak{S}$ is a complete space (vollständiger Raum);† that is to say, if $\{P_1, P_2,$

* Cf., for example, Fréchet, *Sur quelques points du calcul fonctionnel*, Rendiconti del Circolo Matematico di Palermo, vol. 22 (1907), p. 1, and Hausdorff, *Grundzüge der Mengenlehre*, 1914, p. 211.

† Hausdorff, loc. cit., p. 315.

$\cdots, P_n, \cdots\}$ is a "regular" sequence of elements of $\mathfrak{S}$, in other words, if for every $\epsilon > 0$ there exists an integer n_ϵ such that $\overline{P_\lambda P_\mu} < \epsilon$ for $\lambda > n_\epsilon$ and $\mu > n_\epsilon$, there exists a limit element P (i.e., an element P with the property $\lim_{n \to \infty} \overline{P_n P} = 0$).

(3) $\mathfrak{S}$ contains a denumerable subset that is dense in $\mathfrak{S}$.

(4) $\mathfrak{S}$ has no isolated points.

The reader will have no difficulty in introducing in Section 2 the slight modifications that are requisite for making the definitions and the reasoning applicable to a metric, complete space $\mathfrak{S}$ with a dense, denumerable subset and without isolated points. By way of illustration, we show * that the distinction between exhaustible and residual sets may always be made for such a space. For suppose $M_1, M_2, \cdots, M_n, \cdots$ are all non-dense subsets of $\mathfrak{S}$. Let $\epsilon_1, \epsilon_2, \epsilon_3, \cdots$ be a sequence of positive numbers with $\lim \epsilon_n = 0$. Since M_1 is non-dense in $\mathfrak{S}$, its complement $\overline{M}_1 = \mathfrak{S} - M_1$ contains a point A_1 which is in the interior of $\overline{M}_1$ (i.e., which is not a limit point of M_1). Let C_1 be a "sphere" † with center A_1 and radius $< \epsilon_1$ containing only points of $\overline{M}_1$. C_1 contains in turn an interior point A_2 of $\overline{M}_2 = \mathfrak{S} - M_2$. Let C_2 be a sphere with center A_2 and radius $< \epsilon_2$ lying entirely in C_1 and containing only points of $\overline{M}_2$. We thus define the sequence of spheres C_n ($n = 1, 2, \cdots$). The sequence of points $A_1, A_2, \cdots$ is evidently a regular sequence, since $A_\lambda, \lambda > n$, lies in the sphere C_n, hence its distance from A_n is less than ϵ_n, and therefore according to property (1) of $\mathfrak{S}$ the distance between A_λ and A_μ ($\lambda > n, \mu > n$) is less than 2ϵ. Therefore, according to property (2) of $\mathfrak{S}$ there exists a point A of $\mathfrak{S}$ which is the limit of A_n.‡ A lies in every C_n, hence outside of every M_n. Therefore $\mathfrak{S}$ cannot be represented as the sum of a denumerable set of non-dense sets; that is, $\mathfrak{S}$ is not exhaustible. Since the sum of two exhaustible sets is exhaustible, it follows that a residual set, which is the complement in $\mathfrak{S}$ of an exhaustible set, is not itself exhaustible.

Further details of the extension of the ideas of the preceding section to $\mathfrak{S}$ we leave to the reader. We may state the following

THEOREM IV. *Let $\mathfrak{S}$ be any complete, metric space without isolated points and containing a dense, denumerable subset; and $f(P)$ any real function defined for the elements P of $\mathfrak{S}$.§ Then there exists a residual set R such that if P is a point*

* Cf., for example, Hausdorff, loc. cit., chap. VIII, Section 9, especially Theorem VI and p. 328.

† A "sphere" of $\mathfrak{S}$ with center P and radius r is the set of points of $\mathfrak{S}$ whose écart from P is $\leq r$.

‡ It is to be observed that if the radius of C_n does not approach 0, no conclusion with reference to the existence of a limit point of $\{A_n\}$ can be made. For this sequence would then not necessarily form a regular sequence; there may therefore be no limit point unless $\mathfrak{S}$ were assumed to be compact. In this case, however, function space would not be an instance for our results, since a "sphere" of function space, while possessing the properties demanded of $\mathfrak{S}$, is not compact.

§ In the terminology of E. H. Moore, f is a function on $\mathfrak{S}$ to A, where A is the set of real numbers.

of R, and $N_<$ a partial neighborhood of P, the function f is inexhaustibly, and therefore densely, approached at P via $RN_<$. Also there exists a dense subset D of $\mathfrak{S}$ such that $f(P)$ is continuous if P ranges over D.

As particular examples of a complete metric space with a dense denumerable subset and without isolated points, we mention:

(*a*) Euclidean n-space where the écart between two points is the euclidean distance between them.

(*b*) A perfect subset of a euclidean space.

(*c*) Hilbert space, that is, the ensemble $\mathfrak{S}$ of sequences $(x_1, x_2, \cdots, x_n, \cdots)$ of real numbers with convergent $\sum_{n=1}^{\infty} x_n^2$. The écart between two "points" $(x_1, x_2, \cdots, x_n, \cdots)$ and $(y_1, y_2, \cdots, y_n, \cdots)$ is defined to be

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots}.$$

The denumerable set consisting of all the sequences $(r_1, r_2, \cdots, r_n, 0, 0, 0, \cdots)$, where the r's are rational numbers, is dense in $\mathfrak{S}$. Furthermore, $\mathfrak{S}$ is evidently complete. For if $P', P'', \cdots, P^{(n)}, \cdots$ is a regular sequence of points of $\mathfrak{S}$, and $P^{(n)} = (x_1^{(n)}, x_2^{(n)}, \cdots)$, then for $\epsilon > 0$ an integer i exists such that

$$\sqrt{\sum_{n=1}^{n=\infty} (x_n^{(p)} - x_n^{(q)})^2} < \epsilon$$

for $p > i$ and $q > i$. Hence $x_n^{(p)}$ converges to x_n, say; therefore, $P^{(n)}$ to $(x_1, x_2, \cdots, x_n, \cdots)$, which, as may be seen, again belongs to Hilbert space. $\mathfrak{S}$, of course, has no isolated points.

(*d*) Function space: $\mathfrak{S}$ consists of all real continuous functions $f(x)$ defined for $0 \leqq x \leqq 1$. The écart between $f_1(x)$ and $f_2(x)$ is defined to be max $|f_1(x) - f_2(x)|$; the postulates for écart are then satisfied. That there is a dense, denumerable subset follows from the theorem of Weierstrass that every continuous function is the limit of a uniformly convergent sequence of polynomials, which may be assumed to have rational coefficients and are thus denumerable. Moreover, since the limit of a uniformly convergent sequence of continuous functions is continuous, it follows that a regular sequence has a limit, and therefore $\mathfrak{S}$ is complete. $\mathfrak{S}$ obviously has no isolated points.

The assumption that f is single-valued may also be dropped without invalidating Theorems I and II of the last section; Theorem III, of course, implies single-valuedness by its very nature. We thus get the following results:

THEOREM I$^{(a)}$. *Let $f(x, y)$ be any real function defined for the entire XY plane and taking at every point at least one value; the number of values may change, however, from point to point and vary from 1 to c, the cardinal number of the continuum. Then the points (x, y) such that every surface point $(x, y,$*

$f(x, y)$) is densely approached by the surface $z = f(x, y)$ constitute an exhaustible set.

THEOREM II$^{(a)}$. *If $f(x, y)$ is any real, single- or many-valued function defined for the entire XY plane, there exists in the XY plane a residual set R, such that if (x, y) is any point of R, and $N_<$ a partial neighborhood of (x, y), then every point $(x, y, f(x, y))$ is inexhaustibly approached via $RN_<$.*

The proof for these more general results is essentially the same as that for Theorems I and II and is left to the reader.

Theorems I$^{(a)}$ and II$^{(a)}$ hold also for any complete, metric space with a dense, denumerable subset and without isolated points.

3. METRIC PROPERTIES

As in the case of the descriptive properties of Section 1, we shall confine the discussion in this section to planar sets and to one-valued functions of two variables.

Let S be any planar set; P, a point of S; C_r, a circle (interior and boundary) with P as center and r as radius; $m(C_r)$, the area of C_r, and $m_e(SC_r)$, the exterior Lebesgue measure of the portion of S in C_r. Then if

$$\lim_{r \to 0} \frac{m_e(SC_r)}{m(C_r)}$$

exists and is equal to l, we shall say that the *exterior metric density* of S at the point P is l. We have the following

THEOREM V.* *Let S be any planar set. Then the points of S at which the exterior metric density of S is $\neq 1$, i.e., the points where the exterior metric density either does not exist or does exist and is < 1, constitute a set of zero measure (Lebesgue).*

Proof. If P is a point at which the exterior metric density of S does not exist or exists and $\neq 1$, the

$$\liminf_{r \to 0} \frac{m_e(SC_r)}{m(C_r)}$$

must be < 1. Let T_k be the totality of points of S where

$$\liminf_{r \to 0} \frac{m_e(SC_r)}{m(C_r)} < k < 1.$$

The set whose measure we are to prove equal to zero is the sum of $T_{1/2}$, $T_{3/4}$, $\cdots$, $T_{(n-1)/n}$, $\cdots$; and since the sum of $\aleph_0$ sets each of zero measure is again

* For the case where only measurable sets are admitted, cf. Lebesgue, *Leçons sur l'Intégration*, 1904, pp. 124–125; Denjoy, Journal de Mathématiques, ser. 7, vol. 1 (1915), p. 132; Lusin and Sierpinski, Rendiconti del Circolo Matematico di Palermo, vol. 42 (1917), p. 167; de la Vallée Poussin, *Cours d'Analyse*, vol. 2, 1912, p. 114. For the linear case of general (not necessarily measurable) sets, cf. Blumberg, Bulletin of the American Mathematical Society, vol. 25 (1919), p. 350.

of zero measure, it is sufficient to prove that T_k is of zero measure. Let m be the exterior measure of T_k so that T_k may be enclosed, in such a way that every point of T_k is an interior point of at least one C_n, in a sequence of circles C_n, $n = 1, 2, \cdots$, with the sum of their areas $< m + \epsilon$, where ϵ is arbitrarily small. Associate with every point P of T_k a circle C_P lying in the interior of one of the circles C_n, and such that

$$\frac{m_e(C_P S)}{m(C_P)} < k$$

and, a fortiori,

$$\frac{m_e(C_P T_k)}{m(C_P)} < k,$$

where

$$m(C_P) = \text{area of } C_P;$$

this is possible because of the assumed property of the points of T_k. From the circles C_P a denumerable number may be extracted having the same totality of interior points, and from this denumerable number, a finite number C', C'', $\cdots$, $C^{(h)}$, that cover T_k except for a set of exterior measure $< \epsilon$. We now select from the $C^{(n)}$ a subset of non-overlapping circles as follows. Let D' be a $C^{(n)}$ having the maximum area attained by the $C^{(n)}$; remove from the set $\{C^{(n)}\}$ the circles which intersect D', and let D'' be one of the circles left, and having the maximum area for these circles; remove now all circles intersecting D'', and let D''' be one of the circles now remaining and having the maximum area for all circles intersected neither by D' nor by D''. In this way we obtain a set of non-overlapping circles $D^{(n)}$. Let a_1 represent the area of the portion of the plane covered by one or more $C^{(n)}$, and a_2 the portion of the plane covered by the $D^{(n)}$; then it follows from the manner in which the $D^{(n)}$ were chosen that

$$a_2 > \frac{1}{9} a_1.$$

Since every $D^{(n)}$ is a $C^{(n)}$ we have

$$\frac{m_e(D^{(n)} T_k)}{m(D^{(n)})} < k;$$

and since the $D^{(n)}$ do not overlap, we may cover the subset of T_k that lies in the $D^{(n)}$ by means of circles having a total area $< ka_2$. We may thus cover T_k by means of circles having as sum of their areas a number less than $ka_2 + (a_1 - a_2) + \epsilon$, the first term representing a sufficient amount for the portion of T_k in the $D^{(n)}$, the second term, for the portion of T_k in the $C^{(n)}$ and outside of the $D^{(n)}$, and the third term, for the portion of T_k outside of the $C^{(n)}$. Therefore,

$$m < ka_2 + (a_1 - a_2) + \epsilon = a_1 - (1 - k)a_2 + \epsilon;$$

$a_1 - a_2(1-k)a_2$

and in virtue of the inequalities

$$m - \epsilon < a_1 < m + \epsilon, \qquad a_2 > \frac{1}{9} a_1,$$

we have

$$m < m + \epsilon - \frac{1}{9}(1 - k)(m - \epsilon) + \epsilon,$$

whence

$$m < \frac{(19 - k)}{1 - k}\epsilon,$$

and therefore $m = 0$.

Definition. $N_<$ is said to be a *non-vanishing* partial neighborhood of P, if the exterior metric density of $N_<$ at P is $\neq 0$.

We have the following lemma which corresponds to Lemma I in the case of the descriptive properties.

LEMMA II. *Let $\Re$ be a closed relation as in Lemma* I. *The points P for which* (*a*) *$N\Re P$ for every neighborhood of P and* (*b*) *a non-vanishing partial neighborhood $N_<$ exists such that $\overline{N_< \Re P}$ (i.e., $N_< \Re P$ is false) constitute a set of zero measure.*

Proof. Let T be the totality of points P having the properties (*a*) and (*b*). We may then show as in the proof of Lemma I that $\overline{N_< \Re B}$ for every point B of $NN_<$, which, if $N_<$ is a non-vanishing neighborhood of P, is also a non-vanishing neighborhood of P. Every point P of T thus has a non-vanishing neighborhood every point of which does not belong to T; therefore at no point of T is the exterior metric density of T equal to 1; T is therefore of zero measure.

Let $z = f(x, y)$ be a given surface; $P = (x, y)$, a point of the XY plane; and the relation $\Re_{r_1 r_2}$, $r_1 < r_2$ (as in the case of Section 1), such that $I\Re_{r_1 r_2} P$ when and only when there is a point (x, y, ζ), $r_1 \leqq \zeta \leqq r_2$, which is a limit of I'. It follows as before that $\Re_{r_1 r_2}$ is closed. Therefore, by means of Lemma II, we may conclude that the points P for which (*a*) $N\Re_{r_1 r_2} P$ for every neighborhood N of P, and (*b*) $\overline{N\Re_{r_1 r_2} P}$ for some non-vanishing partial neighborhood of P constitute a set $T_{r_1 r_2}$ of zero measure. Let T be the sum of all the sets $T_{r_1 r_2}$, as r_1 and r_2, $r_1 < r_2$, vary over all the rational numbers; T, being the sum of $\aleph_0$ sets of measure zero, is itself of measure zero. Now a point of the surface $z = f(x, y)$ is either isolated (and then (x, y) belongs to a denumerable set of points and hence to a certain set of measure zero) or else it is a limit point of the surface, and then $N\Re_{r_1 r_2} P$ for every neighborhood N of P, if $r_1 < z < r_2$. We therefore have the following

THEOREM VI. *Let $f(x, y)$ be any real, one-valued function defined in the entire plane. Then there exists in the XY plane a set Z, dependent on f, of measure zero, such that if* (*a*) *(x, y) is any point of the XY plane not belonging*

to Z; (b) $N_<$, any non-vanishing partial neighborhood of (x, y); and (c) S, any sphere with $(x, y, f(x, y))$ as center; then there is at least one point of the surface $z = f(x, y)$ lying in the sphere S and having as projection upon the XY plane a point in $N_<$.

That the theorem becomes false if we omit the restriction that the partial neighborhood $N_<$ shall be non-vanishing, is seen from the following example. Let A be a planar non-dense perfect set of positive measure, and B its complement. Let $f(x, y) = 0$ in A and 1 in B. Since A is non-dense and closed, B is an everywhere-dense I-region and therefore every point (x, y) of A has as one of its partial neighborhoods a set $N_<$, consisting exclusively of points of B. For such an $f(x, y)$, (x, y) and $N_<$, however, the assertion of the theorem is false, although, since A is not of zero measure, there are points of A not belonging to the alleged set Z of the theorem. Of course, in virtue of Theorem V, the points of A, with the exception of those belonging to a set of measure zero, have none but vanishing partial neighborhoods that consist entirely of points of B.

It is obvious, as, indeed, the last example shows, that the points at which a function is not densely approached need not be of zero measure. If in passing, however, we alter the meaning of "densely approached" in a rather natural way, we may conclude, according to Theorem VI, that f is densely approached everywhere except at the points of a set of measure zero. The new definition of dense approach is as follows: f is said to be densely approached at P if for every $\epsilon > 0$ and for every *non-vanishing* partial neighborhood $N_<$ of P, there is in $N_<$ a point different from P where f differs from $f(P)$ by less than ϵ.

Let now the relation $\mathfrak{R}_{r_1 r_2}$, $r_1 < r_2$, be such that $I\mathfrak{R}_{r_1 r_2} P$ when and only when there is a point (x, y, ζ), $r_1 \leqq \zeta \leqq r_2$, inexhaustibly approached via I; that is to say, for every $\epsilon > 0$ and for every circle C of the XY plane with P as center, there is an inexhaustible set of points in CI at which f differs from ζ by less than ϵ. Then it follows that $\mathfrak{R}_{r_1 r_2}$ is closed. If f is inexhaustibly approached at P and a non-vanishing partial neighborhood $N_<$ of P exists via which f is not inexhaustibly approached at P, then there exist two rational numbers r_1 and r_2, $r_1 < f(P) < r_2$, such that $N\mathfrak{R}_{r_1 r_2} P$ for every neighborhood N of P and $\overline{N_< \mathfrak{R}_{r_1 r_2} P}$. Therefore according to Lemma II, P belongs to a certain set $Z_{r_1 r_2}$ of measure zero associated with the closed relation $\mathfrak{R}_{r_1 r_2}$. The sum of the $Z_{r_1 r_2}$ for all possible rational numbers $r_1 < r_2$ is also of measure zero; we thus have

THEOREM VII. *The set of points P of the XY plane, at which f is inexhaustibly approached and for which a non-vanishing partial neighborhood exists via which f is exhaustibly approached at P, constitute a set of measure zero.*

It is, of course, not true that the points where f is exhaustibly approached form a set of measure zero. Because of this fact the result for inexhaustible

approach is not as pleasing as that for the case we shall consider next. We shall say that *f is neglectably approached at the point P*, if a sphere S exists with P' as center such that the projection upon the XY plane of the points of the surface that lie in S constitutes a set of zero measure; f is said to be *neglectably approached at the point P via* $N_<$ if a sphere S exists with P' as center such that the projection upon the XY plane of $N'_< S$ has in common with $N_<$ a set of measure zero. We have the following

THEOREM VIII. *Let $z = f(x, y)$ be any real, single-valued function defined in the entire XY plane. Then the points P of the XY plane that possess a non-vanishing partial neighborhood via which f is neglectably approached at P constitute a set of zero measure.*

The proof of this theorem, which is a generalization of Theorem VI, is analogous to that of the latter. A few indications will therefore suffice. We note first that the set of points where f is neglectably approached is of measure zero; the proof is similar to that of the exhaustibility of the set of points at which f is exhaustibly approached.* We then define $I\mathfrak{R}_{r_1 r_2} P$, $P \equiv (x, y)$, as follows: there is a point (x, y, ζ), $r_1 \leqq \zeta \leqq r_2$, such that if S is any sphere with (x, y, ζ) as center, and the surface points in S are projected upon the XY plane, the portion falling in I forms a set of positive exterior measure; i.e., (x, y, ζ) is not neglectably approached via I. It will be seen that $\mathfrak{R}_{r_1 r_2}$ is closed. The rest of the argument is left to the reader.

Theorem VIII asserts that certain sets are of positive exterior measure. We shall go a little further by considering the question of exterior metric density. Let P be any point and $M_{P\epsilon}$ the set of points Q for which

$$|f(Q) - f(P)| < \epsilon.$$

We shall say that f is *quasi-continuous*† at P if for every ϵ the set $M_{P\epsilon}$ has 1 as exterior metric density at P. We have the following theorem, which generalizes Theorem VIII:

THEOREM IX. *f is quasi-continuous except at the points of a set of measure zero.*

Proof. Let $S_{r_1 r_2}$, $r_1 < r_2$, be the set of points (x, y) for which

$$r_1 < f(x, y) < r_2,$$

and $T_{r_1 r_2}$ be the set of points where the exterior metric density of $S_{r_1 r_2}$ is not 1; $T_{r_1 r_2}$ is therefore of measure zero according to Theorem V. If r_1 and r_2 are rational, we have in all $\aleph_0$ sets $T_{r_1 r_2}$, and thus the points in all the $T_{r_1 r_2}$ constitute a set Z of zero measure. If P is a point at which f is not quasi-continuous, then, according to definition, there exists a positive number ϵ such that

* See Section 1.

† Cf. Denjoy, Bulletin de la Société Mathématique de France, vol. 43 (1915), p. 165.

$M_{P\epsilon}$ is not of exterior metric density 1 at P. Let r_1 and r_2 be two rational numbers such that

$$f(P) - \epsilon < r_1 < f(P) < r_2 < f(P) + \epsilon.$$

It follows that P belongs to $T_{r_1 r_2}$ and therefore to Z.

Suppose now that $N_<$ is a partial neighborhood of P such that if C_r is a circle with P as center and r as radius then

$$\liminf_{r \to 0} \frac{m(C_r N_<)}{m(C_r)} = k > 0;$$

the symbol m denotes Lebesgue measure, which $N_<$, as an I-region, possesses. We shall then say that $N_<$ is a "*properly* non-vanishing partial neighborhood" of P. Suppose further that a positive ϵ exists such that

$$\liminf_{r \to 0} \frac{m_e(C_r N_< M_{P\epsilon})}{m(C_r N_<)} = l < 1.$$

We shall then say that the exterior metric density of $M_{P\epsilon}$ is $\neq 1$ at P via $N_<$. It follows then that the surface $z = f(x, y)$ is not quasi-continuous at P. For if δ is any positive number, there exists a circle C_r with arbitrarily small radius and with P as center such that

$$m(C_r N_<) > (k - \delta) m(C_r),$$

and

$$m_e(C_r N_< M_{P\epsilon}) < (l + \delta) m(C_r N_<).$$

Therefore

$$\begin{aligned} m_e(C_r M_{P\epsilon}) &= m_e(C_r N_< M_{P\epsilon}) + m_e(C_r \{C_r - N_<\} M_{P\epsilon}) \\ &< [(l + \delta)(k - \delta) + 1 - (k - \delta)] m(C_r) \\ &= [1 - (1 - l - \delta)(k - \delta)] m(C_r) \end{aligned}$$

for sufficiently small δ; $C_r - N_<$ means here the portion of C_r not in $N_<$. Since $l < 1$, $k > 0$, and δ and r may be made as small as we please, we conclude that the relative exterior measure of $M_{P\epsilon}$ is not 1 at P; that is, that f is not quasi-continuous at P. The totality of such points P is therefore of measure zero. We thus have

THEOREM X. *If P is a point not belonging to a certain set of measure zero, then for every positive ϵ the exterior metric density of $M_{P\epsilon}$ is unity at P via any properly non-vanishing partial neighborhood of P.*

Since an angular region at P, i.e., the portion of the XY plane bounded by two half lines radiating from P, is a properly non-vanishing partial neighborhood of P, it follows as a corollary of Theorem X that, if we neglect a set of measure zero, *the exterior metric density of $M_{P\epsilon}$ is, for every $\epsilon > 0$, equal to unity at P via every angular region with P as vertex.*

Concluding remarks

As in the case of the descriptive properties, it may be seen that the metric theorems may be extended to many-valued functions. Theorem VIII, for example, when thus generalized, would read as follows: *Let $z = f(x, y)$ be any real, single- or many-valued function defined in the entire XY plane. Then the points (x, y) of the XY plane for which a surface point $(x, y, f(x, y))$ and a non-vanishing partial neighborhood $N_<$ exist such that $(x, y, f(x, y))$ is neglectably approached via $N_<$ constitute a set of zero measure.*

It is evident that the metric properties hold for functions of a single variable and, in general, for functions of n variables. Extension to $\aleph_0$-space and to function space would require a satisfactory definition of measure for such spaces;* it is not our purpose to enter upon such questions here. Likewise we shall not attempt to define, by means of postulates, a general space for which our metric results are to hold.

It may be remarked that instead of projecting the surface points of $z = f(x, y)$ upon the XY plane we may project them upon the X-axis and thus obtain other properties. For example, let us define the relationship $\mathfrak{R}_{r_1 r_2 r_3 r_4}$ (between I-regions and points of the X-axis) as follows: $I\mathfrak{R}_{r_1 r_2 r_3 r_4}\xi$ if the surface points having x-coördinates in I have a limit point in the rectangle $x = \xi$, $r_1 \leqq y \leqq r_2$, $r_3 \leqq z \leqq r_4$. $\mathfrak{R}_{r_1 r_2 r_3 r_4}$ is evidently closed. If the r's are taken to be rational, the number of possible relations $\mathfrak{R}_{r_1 r_2 r_3 r_4}$ is $\aleph_0$, and we may, by the aid of Lemma II and an argument repeatedly employed in this paper, obtain the following result: *Let $z = f(x, y)$ be any single- or many-valued function defined in the entire XY plane. Let ξ be a point of the X-axis of the following character: a surface point (ξ, η, ζ) and a partial non-vanishing (linear) neighborhood $N_<$ of ξ exist such that (ξ, η, ζ) is not a limit point of surface points having x-coördinates in $N_<$. The totality of points ξ is of measure zero.*

Similar results may be obtained for other metric properties and also in the case of the descriptive properties. In the case of a function of n variables, we may project upon an $(n - 1)$-space, an $(n - 2)$-space, etc.

* In this connection cf. Gâteaux, Bulletin de la Société Mathématique de France, vol. 47 (1919), p. 47.

University of Illinois,
Urbana, Ill.

A FUNDAMENTAL SYSTEM OF INVARIANTS OF A MODULAR GROUP OF TRANSFORMATIONS*

BY

JOHN SIDNEY TURNER

1. **Introduction.** Let G be any given group of g homogeneous linear transformations on the indeterminates $x_1, \cdots, x_n$, with integral coefficients taken modulo m. Hurwitz† raised the question of the existence of a finite fundamental system of invariants of G in the case where m is a prime p, and obtained an affirmative answer when g is prime to p. Dickson‡ subsequently obtained an affirmative answer for any g.

The general case presents great difficulty, owing to the fact that resolution into irreducible factors with respect to a composite modulus is not, in general, unique. The present investigation is confined to the case in which there are two indeterminates x, y, and m is the square of a prime p. The given group will be denoted by H, the notation G being retained when $m = p$. It is proved that the $p^2 + 1$ invariants

$$L^p, Q^p, pL^\alpha Q^\beta \qquad (\alpha, \beta = 0, 1, \cdots, p-1; \alpha, \beta \text{ not both zero}),$$

where

$$L = yx^p - xy^p, \qquad Q = (x^{p^2-1} - y^{p^2-1})/(x^{p-1} - y^{p-1}),$$

form a fundamental system of (independent) invariants of the group H.

2. Consider the group H of all linear homogeneous transformations modulo p^2:

$$(1) \qquad x' \equiv ax + by, \qquad y' \equiv cx + dy, \qquad ad - bc \equiv 1 \pmod{p^2},$$

where a, b, c, d are integers. To each transformation of H corresponds a unique transformation of the group G:

$$(2) \quad x' \equiv a_1 x + b_1 y, \qquad y' \equiv c_1 x + d_1 y, \qquad a_1 d_1 - b_1 c_1 \equiv 1 \pmod{p},$$

where a_1, b_1, c_1, d_1 are integers. In fact, we have only to choose

$$a_1 \equiv a, \qquad \cdots, \qquad d_1 \equiv d \pmod{p}.$$

Conversely, to each transformation (2) corresponds one or more transformations (1). For, we can choose $a \equiv a_1, \cdots, d \equiv d_1 \pmod{p}$ so that

* Presented to the Society, April 15, 1922.

† Archiv der Mathematik und Physik (3), vol. 5 (1903), p. 25.

‡ *The Madison Colloquium*, Lect. III.

$ad - bc \equiv 1 \pmod{p^2}$. For example, if $a_1 \not\equiv 0 \pmod{p}$ we may take $a \equiv a_1$, $b \equiv b_1$, $c \equiv c_1 \pmod{p}$, and determine d by $ad - bc \equiv 1 \pmod{p^2}$; evidently $d \equiv d_1 \pmod{p}$.

Hence if we reduce all the transformations of H modulo p, we obtain all the transformations of G.

3. **Definition.** A rational and integral invariant of H is a polynomial $I(x, y)$ in x and y with integral coefficients, which remains unchanged modulo p^2 under every transformation (1). That is,

$$I(x', y') \equiv I(ax + by, cx + dy) \equiv I(x, y) \pmod{p^2} \tag{3}$$

for all integers $a, \cdots, d$ such that $ad - bc \equiv 1 \pmod{p^2}$.

Evidently any rational and integral invariant is a sum of homogeneous invariants; hence we restrict the investigation to the latter.

4. Theorem I. *Let $I(x, y)$ be a rational and integral invariant of H, and let $I_1(x, y)$ be the polynomial obtained from $I(x, y)$ by replacing each coefficient by its positive or zero residue modulo p. Then $I_1(x, y)$ will be a rational and integral invariant of G.*

We have (3) for all transformations of H. Now

$$I(x, y) \equiv I_1(x, y) \pmod{p}$$

and

$$I(ax + by, cx + dy) \equiv I_1(ax + by, cx + dy)$$
$$\equiv I_1(a_1 x + b_1 y, c_1 x + d_1 y) \pmod{p},$$

hence

$$I_1(x', y') \equiv I_1(a_1 x + b_1 y, c_1 x + d_1 y) \equiv I_1(x, y) \pmod{p}, \tag{4}$$

and by § 2 this is true for all transformations of G.

5. Now (*Madison Colloquium*, pp. 34–38),

$$I_1(x, y) \equiv kT_1^{\alpha_1} T_2^{\alpha_2} \cdots T_i^{\alpha_i} \cdots T_r^{\alpha_r}* \pmod{p},$$

where k is an integer,

$$T_1 = L, \quad T_2 = Q, \quad T_i = R_i(L^{\frac{1}{2}p(p-1)},\dagger\ Q^{\frac{1}{2}(p+1)}) \quad (i = 3, 4, \cdots, r),$$

R_i being a polynomial in its two arguments, with integral coefficients; moreover the T_i $(i = 1, 2, \cdots, r)$ contain no multiple factors, and are relatively prime modulo p. Hence

$$I(x, y) \equiv kT_1^{\alpha_1} T_2^{\alpha_2} \cdots T_i^{\alpha_i} \cdots T_r^{\alpha_r} + pF(x, y) \pmod{p^2}, \tag{5}$$

where $F(x, y)$ denotes a polynomial in x, y with integral coefficients.

---

* In the discussion which follows, if any α_i is zero the corresponding T_i is to be suppressed.

† If $p = 2$, we omit the divisor 2 in the exponents.

6. Discussion of equation (5). Apply to $I(x, y)$ the transformation

$$x' \equiv x + py, \qquad y' \equiv y \pmod{p^2}, \tag{6}$$

expand by Taylor's Theorem, and denote the partial derivative of T_i with respect to x by T'_i. Then

$$\begin{aligned} I(x+py, y) \equiv I(x, y) &+ pykT_1^{\alpha_1-1} \cdots T_i^{\alpha_i-1} \\ &\times \cdots T_r^{\alpha_r-1} \sum_{i=1}^{r} \alpha_i T_1 \cdots T_{i-1} T'_i T_{i+1} \cdots T_r \pmod{p^2}. \end{aligned} \tag{7}$$

Since (6) is a transformation of H,

$$I(x+py, y) \equiv I(x, y) \pmod{p^2}.$$

Hence either $k \equiv 0 \pmod{p}$, in which case the right member of (5) reduces to its second term, or

$$\sum_{i=1}^{r} \alpha_i T_1 \cdots T_{i-1} T'_i T_{i+1} \cdots T_r \equiv 0 \pmod{p}. \tag{8}$$

Let $(g_i, 1)$ be a point at which $T_i(x, y)$ vanishes. Then, for $j \neq i$, $T_j(x, y)$ cannot vanish at $(g_i, 1)$; for, in that event, $T_j(x, y)$ would be a factor of $T_i(x, y)$ modulo p,* contrary to § 5. Therefore from (8) we have

$$\alpha_i T'_i(g_i, 1) \equiv 0 \pmod{p}. \tag{9}$$

Hence either $\alpha_i \equiv 0$, or $T'_i(g_i, 1) \equiv 0 \pmod{p}$. In the latter case, by a known theorem on Galois imaginaries, $T_i(x, 1)$ and $T'_i(x, 1)$ have a common factor with integral coefficients modulo p. But (§ 5) $T_i(x, 1)$ contains no multiple factor modulo p. Therefore† $T'_i(x, 1) \equiv 0 \pmod{p}$, whence

$$T'_i(x, y) \equiv 0 \pmod{p}. \tag{10}$$

Hence we have

THEOREM II. *In equation* (5), *for each* $i = 1, \cdots, r$, *either* α_i *is a multiple of* p, *or* $T'_i(x, y) \equiv 0 \pmod{p}$.

COROLLARY 1. $\alpha_1 \equiv 0 \pmod{p}$.

For $T_1 = yx^p - xy^p$; hence $T'_1 = pyx^{p-1} - y^p \not\equiv 0 \pmod{p}$.

COROLLARY 2. $\alpha_2 \equiv 0 \pmod{p}$.

For

$$T_2 = x^{p(p-1)} + x^{(p-1)^2} y^{p-1} + \cdots + x^{p-1} y^{(p-1)^2} + y^{p(p-1)},$$

hence $T'_2 \not\equiv 0 \pmod{p}$.

COROLLARY 3. *If* $\alpha_i = p\beta_i$ *for* $i > 2$,

$$T_i^{\alpha_i} \equiv S_i(L^p, Q^p) \pmod{p}, \tag{11}$$

where S_i *is a polynomial in its arguments, with integral coefficients.*

* *The Madison Colloquium*, p. 38.

† Dickson, *Lecture Notes on Double Modulus and Galois Imaginaries*, § 5.

For if we expand

$$T_i^{\alpha_i} = [R_i(L^{\frac{1}{2}p(p-1)}, Q^{\frac{1}{2}(p+1)})]^{p\beta_i},$$

we observe that in each term the exponent of L is a multiple of p and that either the exponent of Q or the coefficient of the term is a multiple of p.

7. **Discussion of $T_i'(x, y) \equiv 0 \pmod{p}$.** Write

$$(12) \qquad T_i(x, y) = \sum_{r=0}^{n} A_r l^{n-r} q^r,$$

where $l = L^{\frac{1}{2}p(p-1)}$,* $q = Q^{\frac{1}{2}(p+1)}$, and the coefficients A_r are integers; then

$$T_i'(x, y) = \sum_{r=0}^{n-1} A_r (n - r) l^{n-r-1} l' q^r + \sum_{r=1}^{n} A_r r l^{n-r} q^{r-1} q' \equiv 0 \pmod{p},$$

where l', q' denote the partial derivatives of l, q with respect to x. Evidently $l' \equiv 0$, $q' \not\equiv 0 \pmod{p}$; therefore

$$(13) \qquad \sum_{r=1}^{n} rA_r l^{n-r} q^{r-1} \equiv 0 \pmod{p}.$$

Each term of (13) is the product of the preceding by cq/l, c a constant; the degree in x of q/l is $\frac{1}{2}(p^2 - p)$,† hence the degrees in x of the successive terms increase by $\frac{1}{2}(p^2 - p)$. Equating coefficients of x, we find in succession

$$nA_n \equiv 0, \quad \cdots, \quad rA_r \equiv 0, \quad \cdots, \quad A_1 \equiv 0 \pmod{p}.$$

Hence in each term of $\sum_{r=1}^{n} A_r l^{n-r} q^r$, either the coefficient A_r or the exponent of q is a multiple of p, and we have

THEOREM III. *If $T_i'(x, y) \equiv 0 \pmod{p}$, then*

$$(14) \qquad T_i(x, y) \equiv S_i(L^p, Q^p) \pmod{p},$$

where S_i denotes a rational and integral function of its arguments, with integral coefficients.

COROLLARY. *$[T_i(x, y)]^{\alpha_i}$ is a polynomial in L^p, Q^p, with integral coefficients, modulo p.*

8. THEOREM IV. *L^p is invariant under the group H.*

Write $L(x', y') = e$, $L(x, y) = f$, where x', y' are derived from x, y by any transformation (2) of the group G; then‡ $e - f \equiv 0 \pmod{p}$. Hence $e - f \equiv 0 \pmod{p}$ for every transformation (1) of H. For if, as in § 2, we choose $a_1 \equiv a, \cdots, d_1 \equiv d$ modulo p, we have

$$L(ax + by, cx + dy) \equiv L(a_1 x + b_1 y, c_1 x + d_1 y) \equiv L(x, y) \pmod{p}.$$

* If $p = 2$, we omit the divisor 2 in the exponents.

† If $p = 2$, the degree is 2.

‡ *The Madison Colloquium*, p. 35.

Also

$$e^p - f^p = (e - f + f)^p - f^p$$
$$= (e-f)[(e-f)^{p-1} + \cdots + \tfrac{1}{2}p(p-1)(e-f)f^{p-2} + pf^{p-1}],$$

and each factor on the right is identically congruent to zero modulo p; hence $e^p - f^p \equiv 0 \pmod{p^2}$; that is

$$[L(ax+by, cx+dy)]^p \equiv [L(x,y)]^p \qquad \pmod{p^2}$$

for every transformation of H.

COROLLARY 1. *In the same way, it can be proved that Q^p is invariant under the group H.*

COROLLARY 2. *$pL^\alpha Q^\beta$ is invariant under the group H.*

9. THEOREM V. *Any rational and integral invariant of the group H is a rational and integral function, with integral coefficients, of the p^2+1 invariants L^p, Q^p, $pL^\alpha Q^\beta$ ($\alpha, \beta = 0, 1, \cdots, p-1$; α, β not both zero). Conversely, any such function is an invariant of H.*

In (5), the term $kT_1^{\alpha_1} T_2^{\alpha_2} \cdots T_r^{\alpha_r}$ is an invariant of H. For if any $\alpha_i \equiv 0 \pmod{p}$, then by Theorems II and IV with their corollaries, $T_i^{\alpha_i}$ is an invariant; $T_1 = L$, $T_2 = Q$, $\alpha_1 \equiv \alpha_2 \equiv 0 \pmod{p}$. While if $\alpha_j \not\equiv 0 \pmod{p}$, $T_j' \equiv 0$, and by Theorems III and IV with their corollaries $T_j^{\alpha_j}$ is an invariant.

Hence the second term $pF(x,y)$ of (5) is an invariant of H, and it follows from § 2 that $F(x, y)$ is an invariant of G. Therefore $pF(x,y)$ is the product of p by a polynomial in L and Q. It follows that if $I(x,y)$ is any rational and integral invariant of H,

$$I(x,y) \equiv S(L^p, Q^p, pL^\alpha Q^\beta) \qquad \pmod{p^2}, \tag{15}$$

where $pL^\alpha Q^\beta$ denotes the set pL, pQ, pL^2, pLQ, $\cdots$, $pL^{p-1}Q^{p-1}$, and S denotes a rational and integral function of its arguments, with integral coefficients.

Conversely, any rational and integral function of L^p, Q^p, $pL^\alpha Q^\beta$, with integral coefficients, is a sum of invariants, and is therefore itself an invariant. Hence these p^2+1 invariants form a fundamental system.

10. THEOREM VI. *The invariants of the fundamental system are independent.*

In view of the coefficients p, neither L^p nor Q^p can be expressed as a polynomial in the remaining invariants, with integral coefficients. Assume that $pL^{\alpha_1}Q^{\beta_1}$, $\alpha_1 \leqq p-1$, $\beta_1 \leqq p-1$, can be so expressed. Then

$$pL^{\alpha_1}Q^{\beta_1} \equiv P(L^p, Q^p, pL^\alpha Q^\beta) \qquad \pmod{p^2} \tag{16}$$

identically in x, y. We may suppose that P contains no group of terms which vanishes identically modulo p^2. Let $mL^{\alpha_2}Q^{\beta_2}$ be any term of P; then $pL^{\alpha_1}Q^{\beta_1}$ and $mL^{\alpha_2}Q^{\beta_2}$ must be of the same total degree in x, y, and also of the same

degree in x alone. Therefore

$$\alpha_1(p+1)+\beta_1 p(p-1)=\alpha_2(p+1)+\beta_2 p(p-1),$$
$$\alpha_1 p+\beta_1 p(p-1)=\alpha_2 p+\beta_2 p(p-1),$$

whence $\alpha_1=\alpha_2$, $\beta_1=\beta_2$. Hence P consists of the single term $pL^{\alpha_1}Q^{\beta_1}$. Evidently $pL^{\alpha_1}Q^{\beta_1}$ is not a product of fundamental invariants, hence the theorem is proved.

11. If we consider the total group

$$(17)\qquad \begin{aligned} &x'\equiv ax+by, \qquad y'\equiv cx+dy && (\bmod p^2),\\ &ad-bc\not\equiv 0 && (\bmod p),\end{aligned}$$

we find, exactly as in Theorem IV, that Q^p is an absolute invariant, and that L^p, $pL^\alpha Q^\beta$ are relative invariants of indices p, α, respectively.

IOWA STATE COLLEGE,
AMES, IOWA.

THE GAUSSIAN LAW OF ERROR FOR ANY NUMBER OF VARIABLES*

BY

J. L. COOLIDGE

The exponential law for the distribution of accidental errors of observation, discovered by Gauss, has been a mathematical classic for over a century. Many have been the attempts to prove it, all based, necessarily, on more or less arbitrary assumptions. Perhaps the most searching examination of it was given by Poincaré in his *Calcul des probabilités*; his final opinion seems to be contained in the following phrase: †

"J'ai plaidé de mon mieux jusqu'ici en faveur de la loi de Gauss dont nous allons maintenant tirer les conséquences. Peut-être pourtant la cause n'était-elle pas parfaitement bonne.

"Elle ne s'obtient pas par des déductions rigoureuses, plus d'une démonstration qu'on a voulu en donner est grossière, entre autres celle qui s'appuie sur l'affirmation que la probabilité des écarts est proportionnelle aux écarts. Tout le monde y croit cependant, me disait un jour M. Lippmann, car les expérimenteurs s'imaginent que c'est un théorème de mathématiques, et les mathématiciens, que c'est un fait expérimental."

The law has been extended to include the distribution of errors depending upon two variables, and in this form it has a certain importance in the theory of ballistics, and in that of statistical correlation; even the case of three variables has been slightly treated. The general case of n variables has never been taken up except in two recent articles by von Mises.‡ The treatment here is based on a very general form of analysis showing how an arbitrary distribution function will lead asymptotically to an exponential form. The analysis is very careful, the point of view extremely abstract, with little relation to practical applications. Moreover, the author gives no indication how the constants should be calculated in any particular case. It is the object of the present paper to deduce the Gaussian law for n variables by a method based upon the classical one for a single variable, but with somewhat broader and more explicit assumptions. In the second part we shall make

* Presented to the Society, December 27, 1922.

† Poincaré, *Calcul des Probabilités*, Paris, 1896, pp. 196 and 149.

‡ *Fundamentalsätze der Wahrscheinlichkeitsrechnung*, Mathematische Zeitschrift, vol. 4 (1919), and *Grundlagen der Wahrscheinlichkeitsrechnung*, ibid., vol. 5 (1920). See also Dodd, *Functions of measurements*, Skandinavisk Aktuarietidskrift, 1922.

the additional assumptions necessary to determine the coefficients in any particular case, and show how these latter may then be calculated in a simple manner.

1. THE DEDUCTION OF THE LAW

Suppose that we are concerned with measuring groups of m quantities. We shall, for simplicity, assume that all groups are equally trustworthy, although the extension to the case of differently weighted groups is not difficult. We shall make certain assumptions about the distribution of errors, meaning, thereby, *accidental errors*, for we assume that constant errors have been removed.

ASSUMPTION 1. *The a priori probability that a group of quantities to be measured should take values in the infinitesimal region*

$$X \pm \tfrac{1}{2}dX, \qquad Y \pm \tfrac{1}{2}dY, \qquad Z \pm \tfrac{1}{2}dZ, \qquad \cdots,$$

where the points $X, Y, Z, \cdots$ lie in a certain continuous m-dimensional manifold S, will differ by an infinitesimal of higher order from the expression

$$f(X, Y, Z, \cdots)\,dX\,dY\,dZ \cdots,$$

where the function f is continuous with continuous first derivatives in S.

ASSUMPTION 2. *The probability that a group of quantities whose true values are $X, Y, Z, \cdots$ in S should be observed to have values, after the removal of constant errors, which lie in the infinitesimal region*

$$x \pm \tfrac{1}{2}dx, \qquad y \pm \tfrac{1}{2}dy, \qquad z \pm \tfrac{1}{2}dz, \qquad \cdots,$$

where $(x, y, z, \cdots)$ is a point of S, will differ by an infinitesimal of higher order from

$$\Phi(X, Y, Z, \cdots, x, y, z, \cdots)\,dxdydz \cdots,$$

where the function Φ is continuous with continuous first and second partial derivatives, and has a value independent of the choice of origin.

The last part of the assumption is plausible in practice, because if we are, for instance, measuring a length on a scale, the accidental errors will arise from various physical causes independent of the position of the 0. Moreover, it has a momentous consequence, for

$$\Phi(X, Y, Z, \cdots, x, y, z, \cdots) = \Phi|0, 0, 0, \cdots, x - X, y - Y, z - Z, \cdots|.$$

Writing in the explicit values of the errors we have

$$\xi = x - X, \qquad \eta = y - Y, \qquad \zeta = z - Z,$$

$$\Phi \equiv \Phi(\xi, \eta, \zeta, \cdots).$$

that is to say, the probability for a system of errors is a function of those errors, and not of the true values and observed values considered as independent variables, a point which has been a stumbling block to some writers.

ASSUMPTION 3. *The mean value for the error on an individual variable is* 0. This again is plausible, for a contrary assumption would show a tendency to favor positive or negative errors, and such a tendency we should naturally class with the constant errors, not with the accidental ones. As a further matter of notation let us write the averages

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}, \qquad \bar{y} = \frac{y_1 + y_2 + \cdots + y_n}{n},$$
$$\bar{z} = \frac{z_1 + z_2 + \cdots + z_n}{n}, \qquad \cdots. \tag{1}$$

The actual errors are

$$\xi_i = x_i - X, \qquad \eta_i = y_i - Y, \qquad \zeta_i = z_i - Z, \qquad \cdots. \tag{2}$$

The residual errors are

$$\delta_i = x_i - \bar{x}, \qquad \epsilon_i = y_i - \bar{y}, \qquad \cdots,$$
$$x_i - \bar{x} = x_i - X - (\bar{x} - X),$$
$$\delta_i = \xi_i - \frac{\sum_j \xi_j}{n}. \tag{3}$$

The individual groups of observations are independent of one another; hence, by Assumption 3,

$$\text{Mean value } \xi_i \xi_j = 0,$$
$$\text{Mean value } \quad \delta_i^2 = \frac{n-1}{n} \text{Mean value } \xi_i^2,$$
$$\text{Mean value } \quad \xi_i^2 = \qquad \text{Mean value } \frac{\sum_j \delta_j^2}{n-1}, \tag{4}$$
$$\text{Mean value } \xi_i \eta_i = \qquad \text{Mean value } \frac{\sum_j \delta_j \epsilon_j}{n-1}.$$

ASSUMPTION 4. *If the infinitesimal increments $dx, dy, dz, \cdots$ be sufficiently small, the probability that the true values lie in the region*

$$\bar{x} \pm \tfrac{1}{2}dx, \qquad \bar{y} \pm \tfrac{1}{2}dy, \qquad \bar{z} \pm \tfrac{1}{2}dz, \quad \cdots$$

is greater than that they lie in any other region of like structure about any other point.

We have now made a sufficient number of assumptions to enable us to deduce the analytic form for our functions. We do this, following the original method of Gauss, by calculating the probability that a given set of observations should have resulted from observing a group of quantities of assumed true value. The probability that the measurements $x_1, y_1, z_1, \cdots, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, \cdots$ were made on quantities whose true values are $X, Y, Z,$

$\cdots$ is by Bayes' theorem

$$\frac{f(X, Y, Z, \cdots)\Phi_1 \Phi_2 \cdots \Phi_n \, dXdYdZ \cdots}{\int \cdots \int f(X, Y, Z, \cdots)\Phi_1 \Phi_2 \cdots \Phi_n \, dXdYdZ \cdots},$$

$$\Phi_i = \Phi(\xi_i, \eta_i, \zeta_i, \cdots).$$

The integration in the denominator is supposed to be extended throughout the whole region S. This expression will be a maximum with the logarithm of its numerator. Equating to 0 the partial derivatives to X, Y, Z, $\cdots$, we get

$$-\frac{\partial \log f}{\partial X} + \frac{\partial \log \Phi_1}{\partial \xi_1} + \frac{\partial \log \Phi_2}{\partial \xi_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \xi_n} = 0,$$

$$-\frac{\partial \log f}{\partial Y} + \frac{\partial \log \Phi_1}{\partial \eta_1} + \frac{\partial \log \Phi_2}{\partial \eta_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \eta_n} = 0.$$

One set of solutions will arise in case all of the observations have been correct, i.e.,

$$-\frac{\partial \log f}{\partial X} + n\frac{\partial \log \Phi(\xi, \eta, \cdots)}{\partial \xi} = 0, \qquad \xi = \eta = \cdots = 0,$$

$$-\frac{\partial \log f}{\partial Y} + n\frac{\partial \log \Phi(\xi, \eta, \cdots)}{\partial \eta} = 0, \qquad \xi = \eta = \cdots = 0.$$

Now, by definition, f is independent of n, hence

$$\frac{\partial \log f}{\partial X} = \frac{\partial \log f}{\partial Y} = \cdots = 0, \qquad f = \text{const.}$$

Let us underline the fact that we are considering probabilities for observations which do not go outside of the region S. We could not have f a constant throughout all space without a contradiction, and it is also evident that f must be rigorously 0 throughout most of space. The partial differential equations now take the simpler form

$$(5) \qquad \begin{aligned} &\frac{\partial \log \Phi_1}{\partial \xi_1} + \frac{\partial \log \Phi_2}{\partial \xi_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \xi_n} = 0, \\ &\frac{\partial \log \Phi_1}{\partial \eta_1} + \frac{\partial \log \Phi_2}{\partial \eta_2} + \cdots + \frac{\partial \log \Phi_n}{\partial \eta_n} = 0, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \, . \end{aligned}$$

These equations hold whenever

$$X = \bar{x}, \qquad Y = \bar{y}, \qquad Z = \bar{z}, \qquad \cdots,$$

$$\xi_1 + \xi_2 + \cdots + \xi_n = 0,$$

$$\eta_1 + \eta_2 + \cdots + \eta_n = 0,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \, .$$

By Assumption 2 we are free to treat our assumed groups as if they were absolutely independent quantities, provided we do not go outside of S. Let us, then, assume that the observed groups $x_1, x_2, \cdots, x_n; y_1, y_2, \cdots, y_n; z_1, z_2, \cdots, z_n$ take such infinitesimal increments that the averages $\bar{x}, \bar{y}, \bar{z}, \cdots$ in (1) are not altered:

$$\frac{\partial}{\partial \xi_1}\left[\frac{\partial \log \Phi_1}{\partial \xi_1}\right]\delta\xi_1 + \frac{\partial}{\partial \xi_2}\left[\frac{\partial \log \Phi_2}{\partial \xi_2}\right]\delta\xi_2 + \cdots + \frac{\partial}{\partial \xi_n}\left[\frac{\partial \log \Phi_n}{\partial \xi_n}\right]\delta\xi_n = 0,$$

$$\frac{\partial}{\partial \xi_1}\left[\frac{\partial \log \Phi_1}{\partial \eta_1}\right]\delta\xi_1 + \frac{\partial}{\partial \xi_2}\left[\frac{\partial \log \Phi_2}{\partial \eta_2}\right]\delta\xi_2 + \cdots + \frac{\partial}{\partial \xi_n}\left[\frac{\partial \log \Phi_n}{\partial \eta_n}\right]\delta\xi_n = 0,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad ,$$

$$\frac{\partial}{\partial \eta_1}\left[\frac{\partial \log \Phi_1}{\partial \eta_1}\right]\delta\eta_1 + \frac{\partial}{\partial \eta_2}\left[\frac{\partial \log \Phi_2}{\partial \eta_2}\right]\delta\eta_2 + \cdots + \frac{\partial}{\partial \eta_n}\left[\frac{\partial \log \Phi_n}{\partial \eta_n}\right]\delta\eta_n = 0.$$

We have also

$$\delta\xi_1 + \delta\xi_2 + \cdots + \delta\xi_n = 0,$$

$$\delta\eta_1 + \delta\eta_2 + \cdots + \delta\eta_n = 0,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot .$$

Each of the first set of equations in the variables $d\xi_1, d\xi_2, \cdots, d\xi_n$ must hold whenever the last equation in these variables holds, hence, integrating once, and dropping subscripts,

$$\frac{\partial \log \Phi}{\partial \xi} = \alpha\xi + \beta\eta + \cdots + \omega.$$

Giving ξ the successive values $\xi_1, \xi_2, \cdots, \xi_n$, and doing the same for the other variables, and then summing, we have

$$\omega = 0,$$

$$\frac{\partial \log \Phi}{\partial \xi} = \alpha\xi + \beta\eta + \cdots,$$

$$\frac{\partial \log \Phi}{\partial \eta} = \alpha'\,\xi + \beta'\,\eta + \cdots, \tag{6}$$

$$\Phi \equiv Re^{-\psi^2(\xi,\eta,\zeta,\cdots)}.$$

Here the expression ψ^2 is a homogeneous quadratic form in the variables.

Our assumptions are sufficient to enable us to make a very definite statement about the function ψ^2, namely, that its discriminant is not zero. For if the discriminant were zero, the partial derivatives would be linearly dependent, and vanish for an infinite number of sets of values for the variables, and this is directly in conflict with our fourth assumption that the only maximum

arose from taking all of these variables equal to zero. Furthermore, since this is a maximum, we know that the form is definite, i.e.,

The homogeneous quadratic form ψ^2 is positive and definite, with a non-vanishing discriminant.

2. Determination of the constants*

It should be emphasized that everything which we have done so far is under the assumption that we are dealing with observations in the region S. We have found the probability that an observation in the region S should lie in a certain infinitesimal sub-region. In practice this is of no interest whatever until we have some idea of what the region S may be. It certainly could not be the whole of space, as the assumption that f is everywhere constant leads to a contradiction. On further consideration we notice two things. First of all, it seems quite plausible that f might be constant throughout a certain region, and equal to 0 almost everywhere else. Second, the expression (6) is excessively small, except in a very strictly confined space, and this rapid diminution of (6) would produce a result close to that of the vanishing of f. In other words, the error in calculating the constants will be very small if we assume that the formula (6) is universally valid. On the strength of this we make

Assumption 5. *For the purpose of calculating constants, formula* (6) *may be assumed true throughout all space.*

We note, secondly, that the only method for calculating our constants is to assume that certain observed values may be identified with their mean values as calculated by formula. The right sides of the last two equations (4) are proportional to the mean values of the averages of certain observed quantities, and we know by Tchebycheff's theorem† that it is highly likely that the value of an average shall be close to its mean value. This leads to

Assumption 6. *When the number of groups is large, the mean values of $\xi^2, \eta^2, \xi, \eta, \cdots$ may be equated to the observed values*

$$\frac{\sum_j \delta_j^2}{n-1}, \quad \frac{\sum_j \epsilon_j^2}{n-1}, \quad \frac{\sum_j \delta_j \epsilon_j}{n-1}, \quad \cdots.$$

These quantities give, of course, the probable errors of individual measure-

* The mathematical manipulation that follows depends on obvious applications of the theory of determinants. The methods and final formula are very close to Greiner, Zeitschrift für Mathematik und Physik, vol. 57, pp. 226 ff., and Pearson, Philosophical Transactions of the Royal Society, vol. 187, pp. 299 ff. Pearson assigns the credit to Edgeworth, Philosophical Magazine, ser. 5, vol. 34, p. 201. I must confess to finding Edgeworth so obscure that I do not know whether his result is like mine or not. Moreover, none of these writers seem to me to set forth the underlying assumptions with desirable clearness.

† Tchebycheff, *Oeuvres*, Petrograd, 1899, vol. 1, p. 687.

ments and their correlation coefficients two by two. It is interesting that these should be the only independent constants.

In order to clarify the manipulation, we shall at this point take the perilous step of changing our notation. Logically, the notation we are now going to adopt might well have been used from the start, but the resulting summation formulas, by their very compactness, are rather obscure, and it is easier to see what is really going on, by using the more diffuse symbolism with many continuation signs which we have employed so far. For the actual errors committed, we shall write

$$(7) \qquad \xi = x_1, \qquad \eta = x_2, \qquad \zeta = x_3, \qquad \cdots,$$

there being m in all. The n sets of m residuals shall be written

$$\delta_{11}, \delta_{12}, \cdots, \delta_{1m}; \quad \delta_{21}, \delta_{22}, \cdots, \delta_{2m}; \quad \cdots; \quad \delta_{n1}, \delta_{n2}, \cdots, \delta_{nm}.$$

Our fundamental formula (6) may now be written

$$(8) \qquad \Phi \equiv Re^{-\Sigma a_{ij}x_ix_j}, \qquad a_{ij} = a_{ji}.$$

The assumptions 5 and 6 may be expressed by the equation

$$(9) \qquad p_{ij} \equiv \frac{\sum_k \delta_{ki}\,\delta_{kj}}{n-1} = R\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_i\,x_j\,e^{-\Sigma a_{ij}x_ix_j}\,dx_1\,dx_2\cdots dx_m.$$

Since the discriminant of our quadratic form is not zero, we may find a linear transformation

$$(10) \qquad x_i = \sum_k c_{ik}\,x'_k, \qquad |c_{ij}| \neq 0,$$

such that

$$\sum_{i,j} a_{ij}\,x_i\,x_j \equiv \sum_{i,j,k,l} c_{ik}\,c_{jl}\,a_{ij}\,x'_k\,x'_l$$

$$= \sum_r b_r\,x'^2_r,$$

$$(11) \qquad \sum_{i,j} c_{ik}\,c_{jl}\,a_{ij} = 0, \qquad k \neq l,$$

$$\sum_{i,j} c_{ir}\,c_{jr}\,a_{ij} = b_r.$$

Since the discriminant is an invariant of weight 2,

$$b_1\,b_2\cdots b_m = |c_{ij}|^2\cdot|a_{ij}|.$$

The inverse of the substitution, contragredient to (10), is

$$w'_k = \sum_i c_{ik}\,w_i,$$

$$w'^2_k = \sum_{i,j} c_{ik}\,c_{jk}\,w_i\,w_j.$$

In the projective space of $n-1$ dimensions where a point has the homogeneous coördinates $x_1, x_2, \cdots, x_n$ the hyperquadric

$$\sum_{i,j} a_{ij}\,x_i\,x_j = 0$$

has the tangential equation

$$\sum_{i,j} A_{ij}\, w_i\, w_j = 0\,, \qquad A_{ij} = \frac{\partial |a_{ij}|}{\partial a_{ij}}\,.$$

In terms of the new variables we have

$$\sum_r b_r\, x_r'^2 = 0\,, \qquad \sum_r \frac{w_r'^2}{b_r} = 0\,,$$

$$|b_1\, b_2 \cdots b_m| \sum_r \frac{w_r'^2}{b_r} = |c_{ij}|^2 \sum_{i,j} A_{ij}\, w_i\, w_j\,, \tag{12}$$

$$|a_{ij}| \sum_r \frac{c_{ir}\, c_{jr}}{b_r} = A_{ij}\,.$$

We may express (9) in terms of the new variables. The only point to remember is that the jacobian of the transformation is $|c_{ij}|$;

$$p_{ij} = R|c_{ij}| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{k,l} c_{ik}\, c_{jl}\, x_k'\, x_l'\, e^{-\sum_r b_r x_r'^2}\, dx_1'\, dx_2' \cdots dx_m'\,.$$

This simplifies greatly because

$$\int_{-\infty}^{\infty} x_k'\, e^{-b_k' x_k'^2}\, dx_k' = 0\,.$$

Hence

$$p_{ij} = R|c_{ij}| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_k c_{ik}\, c_{jk}\, x_k'^2\, e^{-\sum_r b_r x_r'^2}\, dx_1'\, dx_2' \cdots dx_m'\,.$$

We have the well known integrals

$$\int_{-\infty}^{\infty} e^{-b_k x_k'^2}\, dx_k' = \frac{\sqrt{\pi}}{\sqrt{b_k}}\,; \qquad \int_{-\infty}^{\infty} x_k'^2\, e^{-b_k x_k'^2}\, dx_k' = \frac{\sqrt{\pi}}{2 b_k^{3/2}}\,.$$

Hence

$$p_{ij} = R|c_{ij}| \frac{\pi^{m/2}}{2\sqrt{b_1\, b_2 \cdots b_m}} \sum_k \frac{c_{ik}\, c_{jk}}{b_k}\,,$$

or, by (12),

$$p_{ij} = \frac{R}{2} \frac{|c_{ij}|}{|a_{ij}|} \frac{\pi^{m/2}}{\sqrt{b_1\, b_2 \cdots b_m}} A_{ij} = \frac{R}{2} \frac{\pi^{m/2}\, A_{ij}}{|a_{ij}|^{3/2}}\,.$$

Furthermore, since the probability of some group of errors is 1,

$$1 = R \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\sum a_{ij} x_i x_j}\, dx_1 \cdots dx_n\,,$$

$$1 = \frac{R \pi^{m/2} |c_{ij}|}{\sqrt{b_1\, b_2 \cdots b_m}} = \frac{R \pi^{m/2}}{\sqrt{|a_{ij}|}}\,.$$

Dividing out R, we have

$$p_{ij} = \frac{A_{ij}}{2|a_{ij}|} .$$

In these equations the quantities p_{ij} are known; we wish to find the quantities a_{ij}. We first introduce one more symbol:

$$P_{ij} = \frac{\partial |p_{ij}|}{\partial p_{ij}} .$$

Since the process of interchanging each element of a non-vanishing determinant with its cofactor is an involutory one, except for multiplication by a power of the determinant, we must have

$$a_{ij} = MP_{ij} .$$

We calculate M by a little jugglery:

$$|a_{ij}| = M^m |p_{ij}|^{m-1},$$

$$|p_{ij}| = \frac{|A_{ij}|}{2^m |a_{ij}|^m} = \frac{1}{2^m |a_{ij}|},$$

$$M = \frac{1}{2|p_{ij}|} .$$

Let us exhibit our results in the form of a table:
Assumed errors: $x_1, x_2, \cdots, x_n$.
Gaussian law of error for n variables: $\Phi = Re^{-\Sigma a_{ij} x_i x_j}$.
Observed residuals: $\delta_{11}, \delta_{12}, \cdots, \delta_{1m}; \delta_{21}, \delta_{22}, \cdots, \delta_{2m}; \cdots; \delta_{n1}, \delta_{n2}, \cdots, \delta_{nm}$.

$$p_{ij} = \sum_k \frac{\delta_{ki}\,\delta_{kj}}{n-1}; \qquad P_{ij} = \frac{\partial |p_{ij}|}{\partial p_{ij}};$$

$$a_{ij} = \frac{P_{ij}}{2|p_{ij}|}; \qquad R = \frac{1}{\sqrt{(2\pi)^m |p_{ij}|}} .$$

Harvard University,
Cambridge, Mass.

CERTAIN THEOREMS RELATING TO PLANE CONNECTED POINT SETS*

BY

ANNA M. MULLIKIN

I. Introduction

A point set M is said† to be connected if it cannot be expressed as the sum of two mutually exclusive point sets neither of which contains a limit point of the other. Sierpinski‡ has shown that a closed, bounded, connected set of points in space of n dimensions cannot be separated into a countable infinity of closed point sets such that no two of them have a point in common. It will be shown in the present paper that for the case where $n = 2$, this theorem does not remain true if the stipulation that M is closed be removed. It will however be shown that a plane point set, regardless of whether it be closed or bounded, which separates its plane cannot be expressed as the sum of a countable infinity of closed, mutually exclusive point sets, no one of which separates the plane. Of the other results established, the principal one is that if M_1 and M_2 are two closed, connected, bounded point sets, neither of which disconnects a plane S, a necessary and sufficient condition that their sum, M, shall disconnect S is that $\overline{M}$, the set of points common to M_1 and M_2, be not connected.

I wish to thank Professor Robert L. Moore, who suggested the theorems of this paper. Without his help and encouragement it could not have been written.

II

The following is an example of a countable collection of mutually exclusive, closed, and bounded point sets with connected sum. Consider a countable infinity of arcs each of which is made up of four straight-line intervals (Fig. 1), the nth arc being drawn from the point $(m/2^{n-1}, 0)$ to $(m/2^{n-1}, m/2^{n-1})$, thence to $(-m/2^{n-1}, m/2^{n-1})$, thence to $(-m/2^{n-1}, -m/2^{n-1})$ and thence to $(m, -m/2^{n-1})$. Let n go from one to infinity, and let M be the point

* Various parts of this paper were presented to the Society on October 25, 1919, December 28, 1920, and February 26, 1921.

† See N. J. Lennes, *Curves in non-metrical analysis situs with an application in the calculus of variations*, American Journal of Mathematics, vol. 33 (1911), and Bulletin of the American Mathematical Society, vol. 12 (1906).

‡ W. Sierpínski, *Un théorème sur les continus*, Tôhoku Mathematical Journal, vol. 13.

set composed of the sum of all the arcs so obtained. It will be seen that each of these arcs contains a limit point of every subset of M which consists of an infinite number of the remaining arcs. Hence the set M is connected. It is obviously bounded.

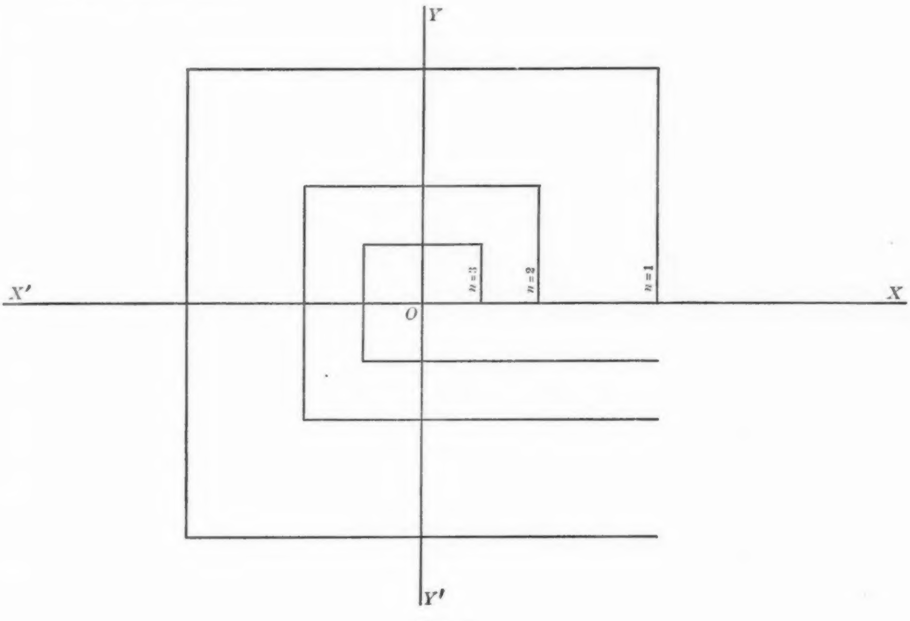

FIG. 1.

THEOREM 1.* *If, in a plane S, K and M are two closed, mutually exclusive point sets and H is a closed, bounded, connected point set having at least one point in common with each of the sets K and M, then there exists a point set $\overline{H}$, a subset of H, such that $\overline{H}$ is connected and contains no point of either K or M, but such that K and M each contain a limit point of $\overline{H}$.*

In our proof of Theorem 1, we shall make use of the following two well known theorems, A and B.

THEOREM A.† *If K and M are two closed point sets having no point in common, and H is a continuous, bounded point set having at least one point in*

* Rosenthal gives a proof for the case in which each of the sets K and M reduces to a single point. See A. Rosenthal, *Teilung der Ebene durch irreduzible Kontinua*, Sitzungsberichte der mathematisch-physikalischen Klasse der Bayerischen Akademie der Wissenschaften zu München, 1919, p. 104.

† Janiszewski gives a proof for the case in which each of the sets K and M reduces to a single point. His proof can readily be extended to the more general case. Cf. S. Janiszewski, *Sur les continus irréductibles entre deux points*, Journal de L'Ecole Polytechnique (2), vol. 16 (1912), p. 109, Theorem 1.

common with each of the sets K and M, then H contains a subset which is irreducibly continuous from K to M.

THEOREM B.* *If $H_1, H_2, H_3, \cdots$ is a countable collection of connected point sets, and P is a point such that every circle containing P contains a point from all except a finite number of these sets, then the limit set† of the sequence $H_1, H_2, H_3, \cdots$ is connected.*

Proof of Theorem 1. By Theorem A, H contains a subset which is irreducibly continuous from K to M. Call this set H_{KM}. Let $\overline{K}$ and $\overline{M}$ denote the points of H_{KM} belonging to K and M respectively. Let H' denote $H_{KM} - \overline{K} - \overline{M}$. We can show that H' is the required $\overline{H}$. Evidently it only remains to be proved that H' is connected. Consider any point P of H'. We can show that the largest connected subset of H' in which P lies has a limit point in either $\overline{K}$ or $\overline{M}$. For suppose it has not. It will then be closed and may be enclosed in a finite number of circles no one of which contains or encloses a point of either $\overline{K}$ or $\overline{M}$. The interiors of these circles form a domain D_P. Now select some point K_1 of $\overline{K}$. Since P and K_1 lie together in the connected set H_{KM} they can be joined by a broken line composed of a finite number of intervals of length less than half an inch, such that the vertices of this broken line belong to H_{KM}. Let L_1 be that vertex on this broken line which immediately precedes the first vertex on it, in the order from P to K_1, that lies without D_P. Then join P and K_1 by a broken line of intervals of length less than a quarter of an inch such that the vertices belong to H_{KM}. Let L_2 be the point on this line corresponding to L_1. Continue this process indefinitely. By Theorem B the limit set will be connected. It will contain P and a point on the boundary of D_P,‡ namely the limit point of $L_1, L_2, L_3, \cdots$. It contains only points of H_{KM}, but since it lies wholly within D_P plus its boundary, it contains no point of $\overline{K}$ or $\overline{M}$ and hence is a subset of H'. This is contrary to the hypothesis that the largest connected subset of H' in which P lies is within the domain D_P.

Denote by H_K the set of those points lying in a connected subset of H' of which $\overline{K}$ contains a limit point, by H_M the set of those which lie in a connected subset of H' of which $\overline{M}$ contains a limit point. Let $H_K + \overline{K}$ be denoted by S_K, $H_M + \overline{M}$ by S_M. Since $S_K + S_M = H_{KM}$, and since H_{KM} is connected, S_K and S_M must have a point in common or else one of these sets must contain a limit point of the other. Suppose first that they have a point in common. This point must belong to H', and it is evident that since H_{KM} is irreducibly

* See S. Janiszewski, loc. cit., p. 98, Theorem 1.

† By the limit set of a sequence of sets $H_1, H_2, H_3, \cdots$ we mean the set of all points $[P]$ such that P is a limit point of a set of points $X_1, X_2, X_3, \cdots$ such that for every k, X_k belongs to H_k.

‡ Janiszewski gives a parallel argument to prove that if the continuous set C contains a point A which is an interior point of the closed set K, then there exists a continuous set containing the point A and contained in K and C. See S. Janiszewski, loc. cit., p. 100, Theorem IV.

continuous, H' must in this case be connected. Suppose secondly that one of the sets contains a limit point of the other, for instance that S_K contains a limit point, P_K, of S_M. And suppose that H' is not connected. Since S_K contains a limit point of S_M it is evident that H_M must exist. Then if $H_M = H'$, H_M cannot be connected. But suppose that the set H_K actually exists; we can show that in this case, too, H_M is not connected. For suppose it were. $P_K + H_M$ together with the largest connected subset of H' in which P_K lies and all points in H_K which are limit points of H_M would then be a connected subset of H' and consequently a proper subset of H'. This together with its limit points in $\overline{K}$ and $\overline{M}$ would be both a continuous set between K and M and a proper subset of H_{KM}. This is contrary to the hypothesis that H_{KM} is irreducibly continuous between K and M. We have therefore shown that if H' is not connected, H_M is not connected.

Suppose this to be the case, and let $H_M = H_{M_1} + H_{M_2}$ where H_{M_1} and H_{M_2} are two mutually exclusive sets neither of which contains a limit point of the other. Suppose P_K is a limit point of H_{M_1}. Then enclose every point of H_{M_2} in a circle which encloses no point of the set $H_{M_1} + P_K + \overline{M}$. The interiors of these circles form a domain D. Now since P_K is a limit point of H_{M_1} and every point of H_{M_1} is connected with some point of $\overline{M}$ in a subset of $H_{M_1} + \overline{M}$, P_K can be joined by an infinite number of broken lines, as before, to points of $\overline{M}$ such that the vertices of these broken lines belong to $H_{M_1} + \overline{M}$, and therefore lie without D or on its boundary. The limit set will then be connected and will contain no point of D. This together with the largest connected subset of H' in which P_K lies and the limit points of this set in $\overline{K}$ will be a continuous set from K to M, a proper subset of H_{KM} since it contains no point of H_{M_2}. This is contrary to the hypothesis that H_{KM} is irreducibly continuous from K to M. We have therefore proved that H' is connected and is the required $\overline{H}$.

THEOREM 2. *If, in a plane S, H is a closed, bounded point set containing two mutually exclusive, closed point sets K and M, but containing no closed, connected subset containing a point of K and a point of M, then it is the sum of two mutually exclusive, closed sets, of which one contains K and the other contains M.*

Proof. There exists a positive number ϵ such that no point of K can be joined to a point of M by a broken line made up of intervals of length less than ϵ such that the end points of these intervals are points of H. For otherwise there would be a closed, connected "limit set" as in Theorem 1. This limit set would belong to H, since H is closed, and it would contain a point of K and a point of M, since K and M are both closed. This is contrary to the hypothesis.

Now let H_1 denote the point set composed of K together with the set of all

points $[P]$ of H such that P can be connected with some point of K by a broken line of intervals of length less than ϵ such that the end points of these intervals belong to H. Let H_2 denote the point set composed of all other points of H. H_2 will contain M and it can easily be seen that neither H_1 nor H_2 contains a limit point of the other, since every point of H_2 is at a distance greater than or equal to ϵ from every point of H_1.

LEMMA. *If M is a closed set not disconnecting* a plane S then any two points of $S - M$ can be joined by a simple continuous arc lying in $S - M$.*

Proof. Let P denote any point of $S - M$. Let S_1 denote the point set composed of P together with all points that can be joined to P by a simple continuous arc lying in $S - M$. Let S_2 denote the set $S - M - S_1$, and suppose that S_2 contains at least one point. Now since M does not disconnect S, S_1 contains a limit point of S_2, or S_2 of S_1. Suppose that S_1 contains a limit point P_1 of S_2. Then enclose P_1 within a square K which neither contains not encloses a point of M. This square will enclose a point P_2 of S_2. Then P_1 and P_2 can be joined by a straight line interval lying within K and therefore containing no point of M. Since P_1 can be joined to P by a simple continuous arc lying in $S - M$ it is obvious that P_2 can also. The argument would be similar in the case where S_2 contains a limit point of S_1. Since either leads to a contradiction we have proved that S_2 does not contain even one point.

THEOREM 3. *If M is the sum of a countable number of closed, mutually exclusive point sets $M_1, M_2, M_3, \cdots$, no one of which disconnects a plane S, then M does not disconnect S.*

Proof. Suppose on the contrary that $S - M = S_1 + S_2$, where S_1 and S_2 are mutually exclusive and neither contains a limit point of the other. Let $\overline{M}_1$ denote a point set composed of those points of M that are limit points of S_1 but not of S_2, $\overline{M}_2$ the point set composed of those that are limit points of S_2, but not of S_1, and let $\overline{M}$ denote the point set composed of those points of M that are limit points of neither S_1 nor S_2. There exists in S a countable collection K_{P_1} of squares $K_1, K_2, K_3, \cdots$ obtained in the following manner. Take any point P_1 of S_1 as the center of a square $\overline{K}_1$ of side 2 inches. Let K_1, K_2, K_3, K_4 be the four squares of side one inch each contained in $\overline{K}_1$ and taken in any order. In general let $\overline{K}_n$ be a square of side 2^n inches which has P_1 for its center and has its sides parallel to those of $\overline{K}_1$, and let it be divided into 2^{4n-2} squares, each of side $1/2^{n-1}$ inches, and let these 2^{4n-2} squares follow each other in any order, and let them be the squares $K_{\frac{2^{4n-2}+11}{15}}, \cdots, K_{\frac{2^{4n+2}-4}{15}}$ in the set K_{P_1}.

Now consider the first square K_1' of K_{P_1} which contains P_1 and satisfies

* M is said to disconnect S if $S - M$ is the sum of two mutually exclusive point sets neither of which contains a limit point of the other.

condition (1) that it contain and enclose only points of $S_1 + \overline{M}_1 + \overline{M}$. There evidently exists one such square, since P_1 is not a limit point of S_2. Add to K_1' the first square K_2' of K_{P_1} that encloses no point of the interior of K_1' and that satisfies condition (1) and also condition (2) that it shall contain or enclose at least one point of S_1, and condition (3) that it shall have an interval in common with K_1'. It is evident that there exists a simple closed curve C_2 which is a subset of $K_1' + K_2'$ and such that the interiors of K_1' and K_2' are subsets of the interior of C_2. In general obtain $C_3, C_4, C_5, \cdots$ in the following manner: C_r shall be obtained by adding to C_{r-1} the first square K_r' of K_{P_1} which encloses no point of the interior of C_{r-1} and which satisfies conditions (1) and (2) and contains an interval in common with C_{r-1}. Then C_r is a simple closed curve which is a subset of $C_{r-1} + K_r'$ and whose interior contains the interiors of C_{r-1} and K_r'. It can easily be shown that provided K_r' exists and C_{r-1} can be obtained in this manner, then C_r can also. For let $A_1 A_2$ be an interval common to C_{r-1} and K_r'. Let $A_1 \overline{B} A_2$ (Fig. 2) be an arc

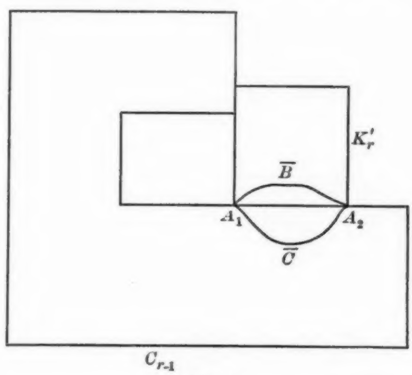

FIG. 2.

lying except for its end points within K_r' and let $A_1 \overline{C} A_2$ be an arc lying except for its end points within C_{r-1}. Then in Theorem 41 of *The foundations of plane analysis situs,** let J_1 be $C_{r-1} - A_1 A_2 + A_1 \overline{B} A_2$ and let J_2 be $K_r' - A_1 A_2 + A_1 \overline{C} A_2$. Then J will be the required C_r. Denote by C the set of curves C_1 $(K_1')\, C_2, C_3, \cdots$.

The sequence C will evidently be infinite unless some C_i contains only points of M. Suppose this to be the case, and suppose first that there is some point P_F of S_1 or S_2 without C_i. The curve C_i divides S into two parts, its interior I, and its exterior E, such that neither of these parts contains a limit point of the other. Now all of C_i must belong to the same set in M,

* See R. L. Moore, *On the foundations of plane analysis situs*, these Transactions, vol. 17 (1916), p. 155.

say M_j, since there does not exist a countable collection of closed, mutually exclusive point sets, consisting of more than one set, whose sum is closed, bounded, and connected.† Then M_j disconnects S; one part of $S - M_j$ is composed of those points of I which do not belong to M_j (P_1 is one such point); the other part is composed of those points of E which do not belong to M_j (by hypothesis there is at least one such point). But this is contrary to the hypothesis of the theorem. Suppose, secondly, that all points of C_i and its exterior belong to M_j. Start from P_1 again to add squares of K_{P_1}, this time adding the extra condition that no square added shall contain or enclose a point of C or its exterior or of M_j; this is possible since M_j and the set composed of C_i together with its exterior are closed sets. Suppose again that some C_k of this new set of curves is a subset of M, and therefore of, say, M_l. If there is some point of S_1 or S_2 without C_k, then, as above, M_l disconnects S, which is contrary to hypothesis. But there must be a point of S_1 or S_2 in the closed, connected, bounded point set made up of $C_i + C_k$ and all points between them, as otherwise this set of points would be the sum of a countable collection of closed, mutually exclusive point sets, subsets of M, and would contain subsets of at least two of the sets $M_1, M_2, M_3, \cdots$, namely of M_j and of M_l. Thus in any case we obtain an infinite sequence C, every curve of which contains at least one point of S_1 and contains only points of $S_1 + \overline{M}_1 + \overline{M}$. Let D_1 denote the sum of the interiors of these curves.

Now suppose there is no point of S_2 without D_1. Then take a point P_2 of S_2 that is within some curve C_{P_2} of C and add squares to it in the manner in which we obtained $K'_1, K'_2, K'_3, \cdots$, except that these squares with their interiors are subsets of $S_2 + \overline{M}_2 + \overline{M}$ and each one contains or encloses at least one point of S_2, and add the extra condition that no one shall contain a point of C_{P_2} or of its exterior. The sum of the interiors of the curves so obtained will be a domain D_2, a subset of the interior of C_{P_2} and therefore not containing all points of S_1. Since D_2 would serve as well for the argument as D_1 we shall suppose that not every point of S_2 is within D_1.

Then since D_1 does not contain all of S it must have some boundary points; let B denote the boundary of D_1. Suppose B contains a point P'_1 of S_1. There is a square $\overline{R}$ of the sequence K_{P_1} which encloses or contains P'_1 but no point or limit point of S_2 and which encloses a point of a curve of the sequence C. Consider the first curve of C which was obtained by adding a square having an interval in common with $\overline{R}$ or its interior. Let R^* be the square so added. Consider two cases.

Case I. It is given that R^* has an interval in common with $\overline{R}$. If the interiors of these squares are mutually exclusive, $\overline{R}$ possesses all of the properties necessary for it to be added in obtaining some curve of C, and it will

† Cf. M. Sierpinski, loc. cit.

subsequently be added or enclosed by the addition of some other square of K_{P_1}. If the interiors of R^* and $\bar{R}$ are not mutually exclusive, evidently the interior of $\bar{R}$ must include that of R^*. But $\bar{R}$ precedes R^* in the sequence K_{P_1} and would have been used instead of R^*, since $\bar{R}$ contains an interval of R^* that R^* has in common with that curve of C to which we supposed it added.

Case II. It is given that R^* has an interval in common with the interior of $\bar{R}$. In this case the interior of R^* lies wholly within $\bar{R}$ but obviously they must also have an interval in common, since R^* was the square added to obtain the first curve of C having an interval in common with $\bar{R}$ or its interior. The argument is the same then as in Case I. We have therefore proved that B contains no point of S_1.

Furthermore B contains no point of S_2. For suppose P_2' is a point of S_2 belonging to B. It is not on any curve of C; therefore it must be a limit point of an infinite number of curves of C. Since P_2' is not a limit point of S_1 there exists a square K_h with P_2' as center, which neither contains nor encloses a point of S_1. Suppose a side of K_h is ϵ_1 inches long. There exists only a finite number of squares of K_{P_1} of side equal to or greater than $\epsilon_1/8$ that have points in common with K_h or its interior. Let Q denote this set of squares. If any square of Q was used in the sequence $K_1', K_2', K_3', \cdots$, let K_n' denote the last square in this sequence that belongs to Q. Then C_n will be the last curve of C formed by adding a square of Q. If no square of Q was so used, let C_n denote any curve of C. Then K_h encloses a point P_l which lies on no curve of the set $C_1, C_2, \cdots, C_n$, but which does lie on a curve of C following C_n and such that the distance from P_2' to P_l is less than $\epsilon_1/4$. Then P_l must be a point on a square of the sequence $K_1', K_2', K_3', \cdots$ of side less than $\epsilon_1/8$ and therefore is at a distance of less than $\epsilon_1/4$ from a point of S_1, since every square of the sequence $K_1', K_2', K_3', \cdots$ contains or encloses a point of S_1. This point of S_1 would then lie within K_h, which leads to a contradiction. We have therefore proved that B is a subset of M.

We can now prove that two closed, mutually exclusive point sets neither of which disconnects S cannot together disconnect S.† For suppose $M = M_1 + M_2$. We have shown above that not every point of S_2 is in D_1 and that no point of S_2 is on B. Let P_0 denote a point of S_2 without D_1. There is a simple continuous arc from P_0 to P_1. Let B_1 denote the first point of B on this arc in the order $P_0\,P_1$. Suppose that M_1 is that one of the sets M_1 and M_2 to which B_1 belongs. Since M_1 is a closed set not disconnecting S, P_0 can, by the lemma, be joined to P_1 by an arc not containing any point of M_1. Let B_2 be the first point of B on this arc‡ in the order $P_0\,P_1$. Then

† Hausdorff gives a proof for the case when one of the sets is bounded. Cf. F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, Veit, 1914, p. 342.

‡ Hereafter in this paper, "arc" and "simple continuous arc" will be considered synonymous terms.

B_2 belongs to M_2. The set $P_0 B_1 + P_0 B_2$ contains as a subset an arc $B_1 B_2$. Let H_1 be a simple closed curve enclosing B_1 but neither containing nor enclosing any point of M_2 (Fig. 3) and containing only one point L_1 of $B_1 B_2$,

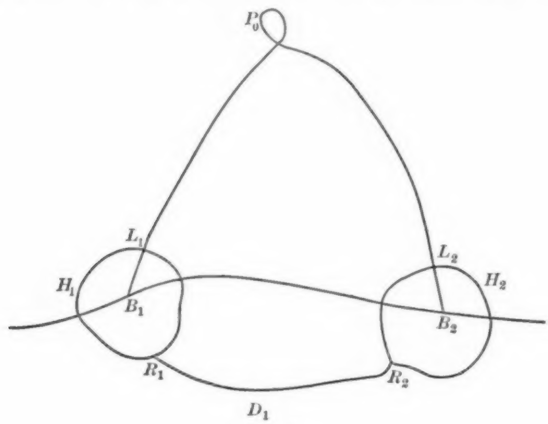

FIG. 3.

and let H_2 be a simple closed curve neither containing nor enclosing any point of H_1 or M_1, but enclosing B_2 and containing only one point L_2 of $B_1 B_2$. Let a point of D_1 within H_1 be joined to a similar point within H_2 by an arc lying within D_1. There will be a subset of this arc, an arc $R_1 R_2$, such that R_1 lies on H_1, R_2 on H_2 and all other points of $R_1 R_2$ lie without both H_1 and H_2. Then there is a simple closed curve J_1 composed of $L_1 L_2 + R_1 R_2$ together with either arc $L_1 R_1$ on H_1 and either arc $L_2 R_2$ on H_2. It is evident that the points of B on or within J_1 that belong to M_1 can be enclosed in a finite number of circles no one of which contains or encloses a point of $L_1 L_2 + L_2 R_2 + R_1 R_2 + M_2$. And similarly those points of B on or within J_1 that belong to M_2 can be enclosed in a finite number of circles no one of which contains or encloses a point of $L_1 L_2 + L_1 R_1 + R_1 R_2 + M_1$ or a point on or within a circle of the first set. Then, clearly, a point on $L_1 L_2$, and therefore without D_1, could be joined to a point on $R_1 R_2$, and therefore within D_1, by an arc lying within J_1 and without both these sets of circles, and therefore not containing a point of B. Since this leads to a contradiction we have shown that if M consists of only two sets it cannot disconnect S. This result can evidently be extended to the case where M consists of any finite number of sets.

Consider the arcs $P_0 B_1$ and $P_0 B_2$ above. Let A_2' denote the first point of $B_2 P_0$ on $P_0 B_1$. If A_2' is different from P_0 it is evident that a point on $B_2 A_2'$ very near A_2' can be joined to P_0 by an arc lying without D_1 and con-

taining only P_0 in common with $P_0 B_1$. From this together with $B_2 A_2'$ we obtain an arc $P_0 B_2$ which lies, except for B_2, without D_1 and has only P_0 in common with $P_0 B_1$. Let M_{B_1} denote that set of M to which B_1 belongs, M_{B_2} the set to which B_2 belongs. Since $M_{B_1} + M_{B_2}$ does not disconnect S, P_0 can be joined to P_1 by an arc not containing any point of M_{B_1} or M_{B_2}. Let B_3 denote the first point of B on this arc in the order $P_0 P_1$, and let M_{B_3} denote that set of the collection M to which B_3 belongs. As above, let A_3' denote the first point $B_3 P_0$ has on the sum of the arcs $P_0 B_1$ and $P_0 B_2$. By taking a point on $B_3 A_3'$ very near A_3', and drawing a suitable arc to P_0, we obtain an arc $P_0 B_3$ which has only P_0 in common with $P_0 B_1$ or $P_0 B_2$, and which lies, except for B_3, without D_1. Similarly obtain the arc $P_0 B_4$ where B_4 belongs to B and to M_{B_4}, such that $P_0 B_4$ has only P_0 in common with $P_0 B_1$, $P_0 B_2$, or $P_0 B_3$, and lies, except for B_4, without D_1. Now it is evident that the sum of two of these arcs is an arc crossing the sum of the other two. Suppose for instance that the arc $P_0 B_1 + P_0 B_3$ crosses the arc $P_0 B_2 + P_0 B_4$. Let H_3 be a simple closed curve (Fig. 4) enclosing B_1 but

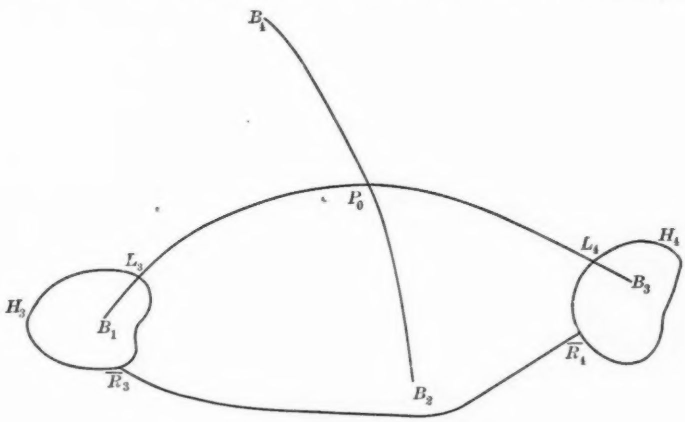

FIG. 4.

neither containing nor enclosing any point of $P_0 B_2$, $P_0 B_3$, $P_0 B_4$, M_{B_2}, M_{B_3}, or M_{B_4} and such that it contains only one point L_3 of the arc $P_0 B_1$. Let H_4 be a simple closed curve enclosing B_3, containing only one point L_4 of the arc $P_0 B_3$, and neither containing nor enclosing any point of H_3, $P_0 B_1$, $P_0 B_2$, $P_0 B_4$, M_{B_1}, M_{B_2}. or M_{B_4}. Now consider the first curve of C, C_A, which has points within both H_3 and H_4. There is an arc $R_3 R_4$, a subset of C_A, such that R_3 is on H_3, R_4 is on H_4 and all other points of $R_3 R_4$ lie without both H_3 and H_4. Select one of the arcs $L_3 R_3$ on H_3 and one of the arcs $L_4 R_4$ on H_4 and let those selected be denoted by $L_3 R_3$ and $L_4 R_4$ throughout

the discussion. Then $L_3 L_4 + L_4 R_4 + R_3 R_4 + L_3 R_3$ is a simple closed curve J_2. Now let $\overline{R}_3 \overline{R}_4$ be an arc on C_A lying on J_2 or within J_2 except for its end points, and having $\overline{R}_3$ on H_3 and $\overline{R}_4$ on H_4 and all other points without both H_3 and H_4, such that if we consider the simple closed curve $\overline{J}_2$ which $\overline{R}_3 \overline{R}_4$ forms with that arc of J_2 (from $\overline{R}_3$ to $\overline{R}_4$) that contains $L_3 L_4$, there is no arc of C_A (except $\overline{R}_3 \overline{R}_4$) lying on $\overline{J}_2$ or within $\overline{J}_2$ except for its end points, and having one point on H_3 and one on H_4. It is possible that the arc $\overline{R}_3 \overline{R}_4$ will be the arc $R_3 R_4$. Now it is evident that either $\smile{P_0 B_2}$* or $\smile{P_0 B_4}$ will lie within $\overline{J}_2$. Suppose $\smile{P_0 B_2}$ does. Then B_2 will lie on or within $\overline{J}_2$. The point B_2 is a limit point of an infinite sequence of curves of C following C_A. Let $\overline{P}_1, \overline{P}_2, \overline{P}_3, \cdots$ denote a sequence of points on successive curves of this sequence such that B_2 is a sequential limit point of the set $\overline{P}_1, \overline{P}_2, \overline{P}_3, \cdots$. If B_2 is on $\overline{J}_2$ these points may all coincide with B_2; if not, let them be chosen so that they lie within $\overline{J}_2$.

Denote by $\overline{C}_i$ that curve of C on which $\overline{P}_i$ lies, and let $\overline{N}_i$ be the last point of $\overline{C}_i$ starting from $\overline{P}_i$ in either order around $\overline{C}_i$ such that $\overline{P}_i \overline{N}_i$ lies on or within $\overline{J}_2$. It will be seen that every $\overline{N}_i$ will have to be on $L_3 \overline{R}_3$ or $L_4 \overline{R}_4$, for if any $\overline{N}_i$ is on $\smile{\overline{R}_3 \overline{R}_4}$ there must be a point F of $\overline{C}_i$ very near $\overline{N}_i$, without $\overline{J}_2$ and therefore on the opposite side of C_A from $\smile{P_0 B_2}$; for if B_2 is not on C_A it can be joined, because of the condition put upon $\overline{R}_3 \overline{R}_4$, to any point P on $\overline{R}_3 \overline{R}_4$ by an arc having only the point P on C_A. Now since $\smile{P_0 B_2}$ is without C_A, F must be within C_A. This is impossible, since F is on a curve following C_A in C, and is therefore on or without C_A. Now the set of arcs $\overline{P}_i \overline{N}_i$ (where $i = 1, 2, 3, \cdots$) determines a limit set Y such that Y is a closed, connected set, every point of which is on or within $\overline{J}_2$. Since B_2 belongs to Y, Y must be a subset of M_{B_2}. But there will be a point of Y on either $L_3 \overline{R}_3$ or $L_4 \overline{R}_4$, namely a limit point of the set of points $\overline{N}_1, \overline{N}_2, \overline{N}_3, \cdots$. This leads to a contradiction, for neither H_3 nor H_4 contains a point of M_{B_2}. The supposition that M disconnects S is therefore proved false.

THEOREM 4. *If M_1 and M_2 are two closed, connected, bounded point sets, neither of which disconnects a plane S, a necessary and sufficient condition that their sum, M, shall disconnect S is that $\overline{M}$, the set of points common to M_1 and M_2, be not connected.*

Proof. The condition is necessary. For suppose that $\overline{M}$ is connected; we can prove that M does not disconnect S. For suppose $S - M = S_1 + S_2$ where S_1 contains no limit point of S_2 nor S_2 of S_1. Let $\overline{S}$ be a square enclosing M. Then it is evident that either S_1 or S_2 must lie within $\overline{S}$. Suppose S_1 does. Let P_1 be a point of S_1, and consider all points which lie

* If $P_0 B_2$ is an arc, $\smile{P_0 B_2}$ denotes the point set $P_0 B_2 - P_0 - B_2$.

with P_1 in a connected subset of S_1. It will be seen that since M is closed, these points form a domain D_s, a subset of S_1, and of the interior of $\overline{S}$, such that the boundary of D_s is a subset of M. Let M' denote the point set $M_1 - \overline{M}_2$ and M'' the point set $M_2 - \overline{M}$. Let P_2 denote a point of S_2 that is without $\overline{S}$; by the lemma, P_2 can be joined to P_1 by an arc not containing any point of M_2. Let P' denote the first point the arc $P_2\,P_1$ has on

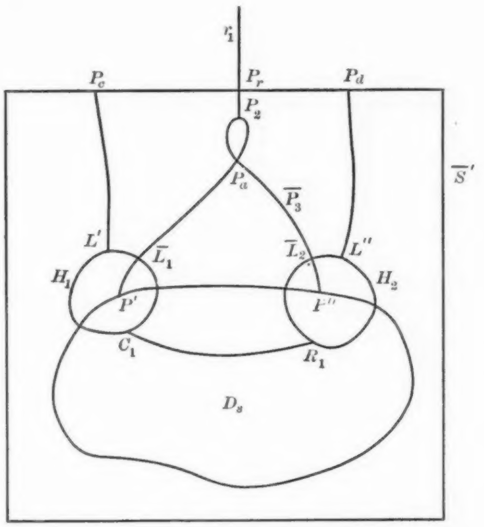

FIG. 5.

the boundary of D_s (Fig. 5), let P_2 be joined to P_1 by an arc not containing any point of M_1, and let P'' denote the first point of this arc $P_2\,P_1$ on the boundary of D_s. Then P' belongs to M' and P'' to M''. Then there exists an arc $P'\overline{P}_3\,P''$, a subset of $P_2\,P' + P_2\,P''$, which lies, except for its end points, without D_s. As in Theorem 3, there exists a simple closed curve H_1 enclosing P', containing only one point $\overline{L}_1$ of the arc $P'\overline{P}_3\,P''$, and neither containing nor enclosing a point of M_2, and there exists a corresponding curve H_2 enclosing P'', containing only one point $\overline{L}_2$ of the arc $P'\overline{P}_3\,P''$ and neither containing nor enclosing a point of M_1 or H_1, and there exists an arc $C_1\,R_1$ in D_s having only C_1 on H_1 and R_1 on H_2. Then there is a simple closed curve $\overline{J}_1$, composed of $\overline{L}_1\,\overline{P}_3\,\overline{L}_2 + C_1\,R_1$ together with the arcs $\overline{L}_1\,C_1$ on H_1 and $\overline{L}_2\,R_1$ on H_2 so chosen that $\overline{J}_1$ encloses P' and P''.

We can show as in Theorem 3 that if those points of the boundary of D_s that this curve contains or encloses belong to $M' + M''$ then two points, one without D_s and the other within D_s, can be joined by an arc lying within

$\overline{J}_1$ and containing no point of the boundary of D_s. Since this is impossible $\overline{J}_1$ must enclose some point of $\overline{M}$, and since $\overline{M}$ is connected and since no point of $\overline{M}$ is on $\overline{J}_1$, $\overline{M}$ must lie wholly within $\overline{J}_1$.

A ray of a straight line can be drawn from P_2, lying wholly without $\overline{S}$, and there exists a ray r_1, a subset of the first ray together with one of the arcs $P_2 P_1$, such that r_1 lies without $\overline{J}_1$ except for its end point P_a on $P'\overline{P}_3 P''$. Let $\overline{S}'$ denote a square enclosing $\overline{S}$ and both arcs $P_1 P_2$ and containing only one point P_r of r_1. If L' is a point on H_1 and $\overline{J}_1$ sufficiently near $\overline{L}_1$ there exists an arc $P_c L'$ such that P_c is a point on $\overline{S}'$ and such that $P_c L'$ lies without D_s, $P_c L' - L'$ lies without $\overline{J}_1$ and $P_c L' - P_c$ lies within $\overline{S}'$. Similarly a point L'' of H_2 and $\overline{J}_1$ can be joined by an arc to the point P_D of $\overline{S}'$, and in such a way that the arc $P_D L''$ contains no point of the arc $P_c L'$. Let $L' L''$ denote the arc composed of $P_c L' + P_D L''$ together with that arc of $\overline{S}'$, from P_c to P_D, which does not contain P_r. The arc $L' L''$ together with the arc $L' C_1 R_1 L''$ of $\overline{J}_1$ makes a simple closed curve $\overline{J}_2$ whose interior has no point in common with that of $\overline{J}_1$. Since $\overline{J}_2$ encloses points of D_s we can show as above that it must contain all of $\overline{M}$. Since this is obviously impossible, the supposition that $\overline{M}$ was connected is proved to be false.

The condition is sufficient. For suppsoe $\overline{M}$ is not connected. Then it is the sum of two closed, mutually exclusive point sets, $\overline{M}_1$ and $\overline{M}_2$. We can show that in this case $S - M = S_1 + S_2$ where S_1 and S_2 are mutually exclusive and neither contains a limit point of the other.

The point set $\overline{M}_1$ contains some point $\overline{P}_1$ which is a limit point of M'' since M_2 is connected and $\overline{M}_1$ and $\overline{M}_2$ are closed point sets. Let J_{p_1} denote a circle enclosing $\overline{P}_1$ but neither containing nor enclosing any point of $\overline{M}_2$. This circle contains a point P_1^* of $M_2 - \overline{M}_1 - \overline{M}_2$, and there is a ray of an open curve from P_1^* not containing any point of M_1, since M_1 is closed and bounded and does not disconnect S. And a subset of this ray is a ray r_2, having an end point on J_{p_1} and lying, except for this end point, without J_{p_1}. Now, by the Heine-Borel theorem, $\overline{M}_2$ can be enclosed in a finite set of circles no one of which contains or encloses any point of $\overline{M}_1$ or J_{p_1} or r_2. Consider any one of these circles together with all of the set that are connected with it. By Theorem 42 of *The foundations*† there exists a simple closed curve J_{c_2}, a subset of these circles, whose interior contains all of their interiors. It will not contain any point of $\overline{M}$, but will enclose some points of $\overline{M}$; let M_2^* denote the set of these points. The curve J_{p_1} with r_2 will be without J_{c_2}. Now, as above, there exists a ray r_3 which has its end point on J_{c_2}, lies, except for this point, without J_{c_2} and contains no point of M_1. Then all points of $\overline{M}$ that are without J_{c_2} can be enclosed in a finite set of circles no one of which contains or encloses any point of J_{c_2} or r_3. If these circles do not form a connected

† Cf. R. L. Moore, loc. cit., p. 156.

point set they can be joined by a finite number of arcs not containing any point of $J_{c_2} + r_3$. For J_{c_2} with its interior does not disconnect S, and r_3 does not disconnect S and they have only one point in common; therefore, by the first part of this theorem, J_{c_2} with its interior and r_3 does not disconnect S. These arcs can be covered by a finite set of circles not containing any point of J_{c_2} or r_3, and hence from all these circles a simple closed curve J_{c_1} can be obtained which has the following properties: it is wholly without J_{c_2} and does not enclose it; it encloses all points M_1^*, of $\overline{M}$, that J_{c_2} does not enclose; therefore it contains no point of $\overline{M}$.

The points of M_2 that are on J_{c_1} or J_{c_2} or without both J_{c_1} and J_{c_2} form a closed set, a subset of M''. This set can be covered by a finite set of circles, T, such that no circle of T contains or encloses any point of M_1. Now there is a ray of an open curve, not containing any point of M_1, from some point within each of these circles. Let $[r_t]$ denote the sum of such rays. And let M_1 be covered by a finite set of circles no one of which contains or encloses any point of the circles T, or of the point set $[r_t] + r_2 + r_3$. Since M_1 is connected this set of circles will be connected, and, as above, there exists a simple closed curve J_1 which is a subset of them and encloses all of their interiors and therefore all of M_1. The curve J_1 cannot wholly enclose either J_{c_1} or J_{c_2} since it contains no point of r_2 or r_3.

Let S^* be a square enclosing J_{c_1}; J_{c_2}, J_1 and M_2. It is evident that there is an arc $P_3 P_4$ composed of a finite number of straight line intervals lying within S^*, except for the points P_3 and P_4 which are on S^*, and which separates the interior of S^* into two parts such that J_{c_1} lies wholly within one part and J_{c_2} within the other. Let A be a point of $\overline{M}$ within J_{c_1} and B a point of $\overline{M}$ within J_{c_2}. By Theorem 1 of this paper there is a connected subset of M_2 lying within J_1 such that A and some point A' of J_1 are limit points of it. Moreover this set will lie within J_{c_1} since it is connected, A lies within J_{c_1} and J_1 encloses no point of M_2 that is on J_{c_1}. Similarly there exists a connected subset of M_2 lying within J_1 and J_{c_2} such that B and some point B' on J_1 are limit points of it.

Now there is a subset, CD, of $P_3 P_4$ which satisfies the following conditions: it is an arc lying except for C and D within J_1; C and D are on J_1 and separate A' and B' on J_1. For consider all arcs of $P_3 P_4$ which lie except for their end points within J_1 and have their end points on J_1. There are a finite number of these, since J_1 is composed of a finite number of arcs of circles and $P_3 P_4$ of a finite number of straight line intervals. Let $L_1, L_2, L_3, \cdots$, denote the set of all such arcs and let L denote the point set $L_1 + L_2 + L_3 \cdots$. Suppose that no L_j separates A' and B' on J_1. Denote by a' and a'' the two arcs $A'B'$ on J_1. If any arc L_i of L has an end point on a' for instance let X_1 denote the first such end point on a' in the order $A'B'$ and X_2 the other

end point of this arc. Then consider the simple closed curve $X_1A' + a'' + B'X_2 + X_2X_1$. It will be seen that A' and B' can be joined by an arc lying on or within this curve which contains no point of L_i. By a finite number of repetitions of this process it is evident that we can obtain a simple closed curve C_x, a subset of J_1 plus its interior, containing A' and B' but enclosing no point of L. A' and B' could then be joined by an arc lying on or within C_x and containing no point of L and hence no point of P_3P_4. Since this is impossible there must exist an arc CD, having the properties stated above.

Since $\overset{\frown}{CD}$, being a subset of P_3P_4, lies without both J_{c_1} and J_{c_2} and since it lies within J_1, it contains no point of M_2. Therefore of the two regions into which it divides the interior of J_1,* one must contain A, the other B. Let R denote the interior of J_1. Let $F_1F_2F_3$ be an arc as indicated in Fig. 6

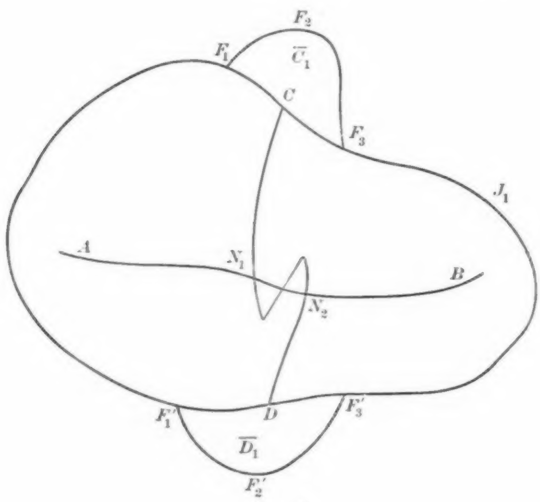

FIG. 6.

with its end points on J_1 and lying, except for its end points, without J_1 and wholly without J_{c_1} and J_{c_2} and such that the simple closed curve $F_1CF_3F_2F_1$ neither contains nor encloses a point of M_1 or M_2. Let $\overline{C}_1$ be a point within this curve. Let $F_1'F_2'F_3'$ be a similar arc forming part of a simple closed curve $F_1'DF_3'F_2'F_1'$ and let $\overline{D}_1$ be a point within it. Now if we suppose that M does not disconnect S, let N_1 be the first point of the arc CD on M_1. Then by the first part of this theorem, $CN_1 + M$ does not disconnect S. Let N_2 be the first point the arc DC has on M_1 (N_1 and N_2 must exist since M_1 is con-

* Cf. R. L. Moore, loc. cit., p. 141.

nected and CD divides R into two parts, of which one contains A, and the other contains B). By several applications of the first part of this theorem, we can show that $M + CN_1 + DN_2 + F_1 CF_3 + F_1' DF_3'$ does not disconnect S. Let H denote this point set. Then $\overline{C}_1$ and $\overline{D}_1$ can be joined by a broken line containing no point of H. By extending this broken line we get a broken line CED which has only C and D on H. By Theorem 43 of *The foundations*† there is a simple closed curve $\overline{C}$, a subset of J_1 together with the broken line CED whose interior is a subset of R, which encloses N_1 but no point of the arc CED. Then $\overline{C}$ will enclose all of $M_1 + (CN_1 - C) + (DN_2 - D)$, since this is a connected point set and has no point on $\overline{C}$. The curve $\overline{C}$ will contain $F_1 CF_3$, $F_1' DF_3'$, A' and B' since it encloses A and B and contains no point of the previously described connected sets of M_2 between A and A', and B and B'. Then it is evident that there is a curve $\overline{C}'$, which has all the above mentioned properties of $\overline{C}$ except that it contains only the points C and D of the broken line CED and which is such that (1) it is a subset of $\overline{C}$ plus its interior, (2) those points of $\overline{C}'$ that are not on $\overline{C}$ do not belong to M.

It is possible that the broken line segment $CD - C - D$ (a subset of $P_3 P_4$) will not lie within $\overline{C}'$. Consider the arcs of the broken line CD which lie except for their end points without $\overline{C}'$ and have their end points on $\overline{C}'$. Let $C'D'$ denote such an arc. C' and D' are points of R. Let E' denote a point of $C' D'$ without $\overline{C}'$. The points C' and D' can be joined by an arc $C' F' D'$ such that $C' F' D'$ lies within $\overline{C}'$ and therefore within J_1. Then the simple closed curve $C' E' D' F' C'$ is a subset of R, and therefore the arc $C' GD'$ of $\overline{C}'$ that it encloses must be a subset of R, and therefore it must belong to that part of $\overline{C}'$ which does not belong to $\overline{C}$ and which therefore contains no points of M. Now there are only a finite number of arcs such as $C' E' D'$, that are parts of the broken line CD, which lie except for their end points without $\overline{C}'$ and have their end points on $\overline{C}'$, and we see that each one can be replaced by an arc lying on $\overline{C}'$ and containing no point of M_2. This gives us a continuous curve from C to D lying on or within $\overline{C}'$ and containing no point of M_2, and there is an arc, a subset of this curve, having the same properties. It is evident that this arc can be replaced by an arc $CE^* D$ lying within $\overline{C}'$ except for C and D and containing no point of M_2. Now C and D separate A' and B' on $\overline{C}'$. For if they did not, A' and B' could be joined by an arc a_1 lying, except for A' and B', within $\overline{C}'$, and C and D could be joined by a similar arc a_2 such that a_2 had no point in common with a_1. But both these arcs would lie except for their end points within $\overline{C}$ and J_1 and have their end points on $\overline{C}$ and J_1. This leads to a contradiction, since C and D separate A' and B' on J_1.

Now the arc CED, together with the arc $CE^* D$, forms a simple closed

† See R. L. Moore, loc. cit., p. 157.

curve J_3 containing no point of M_2, but such that one of the points A and B is within it, and the other without it. This is impossible since M_2 is a connected point set. Therefore the supposition that $\overline{C}_1$ and $\overline{D}_1$ can be joined by an arc containing no point of H is false and since H disconnects S, M must disconnect S. Evidently one of the two sets into which M separates S is the point set, S_1, composed of $\overline{C}_1$ together with all points that can be joined to it by arcs containing no point of M. The set S_2 will then be the point set $S - M - S_1$.

THEOREM 5.* *If M_1 and M_2 are two closed, bounded, connected point sets in a plane S, such that neither M_1 nor M_2 disconnects S and such that M_1 and M_2 have in common only K_1 and K_2, where K_1 and K_2 are mutually exclusive connected sets, then $S - M_1 - M_2$ is the sum of just two mutually exclusive, connected domains.*

Proof. We have shown above that under the conditions of this theorem $S - M_1 - M_2$ is not connected. Suppose then that it is the sum of more than two mutually exclusive connected domains. There will exist three points, P_1, P_2 and P_3, no two of which can be joined by an arc containing no point of $M_1 + M_2$. It is evident that in the preceding theorem the curve J_1 could have been constructed in such a way that P_1, P_2 and P_3 were without it; for there exist three open curves, containing P_1, P_2 and P_3 but containing no point of M_1, and J_1 could have been drawn so as not to contain any point of these open curves. We shall suppose that J_1 has been so drawn. We can furthermore suppose that J_1 is replaced by a polygon W, satisfying the conditions which J_1 satisfies.

Now P_1 and P_2 can be joined by an arc made up of a finite number of straight line intervals containing no point of M_2, since M_2 does not disconnect S; and since $M_2 + P_1 P_2$ does not disconnect S, it is obvious that P_2 and P_3 can be joined by a similar arc which contains no point of M_2 and has only P_2 in common with $P_1 P_2$. Similarly there is a broken line, an arc $P_3 P_1$, containing no point of M_2, and having only P_3 and P_1 in common with $P_1 P_2 + P_2 P_3$. Now any one of these three arcs will contain only a finite number of intervals lying except for their end points in W and having their end points on W. These intervals will be of two kinds, those whose end points separate A' and B' on W, and those whose end points do not separate A' and B' on W. Consider an interval $\overline{X}_1 \overline{X}_2$ of the second sort. Let $\overline{X}_1 X' \overline{X}_2$ denote that arc on W which does not contain A' or B'. Then obviously $\overline{X}_1 \overline{X}_2 + \overline{X}_1 X' \overline{X}_2$ is a simple closed curve enclosing no point of K_1 or K_2, and by the method employed in Theorem 3, we can draw an arc $\overline{X}_1 \overline{X}_2$ such that $\underbrace{\overline{X}_1 \overline{X}_2}$ is within this curve and contains no point of either

* Rosenthal gives a proof for the case in which each of the sets K_1 and K_2 reduces to a single point. See A. Rosenthal, loc. cit., p. 102, Theorem 6.

M_1 or M_2. Let the original interval $\overline{X}_1\,\overline{X}_2$ be replaced by this arc $\overline{X_1\,X_2}$ and let this process be carried out for every such interval, of the second kind, on each of the three arcs. Now each one of the three must have at least one interval of the first kind; for since no one of these arcs has any point of M_2 on it and all of M_1 lies within W and every interval of the second sort of the arcs lying within W has been replaced by an arc containing no point of either M_1 or M_2, if one of these arcs, $P_1\,P_2$, for instance, had no interval of the first sort on it, it would be replaced by an arc $P_1\,P_2$ which had no point of M_1 or M_2 on it, which is contrary to our supposition. Let $Y_1\,Y_2$ be the first interval on $P_1\,P_2$ in the order $P_1\,P_2$ which separates A' and B' on W, let $Y_3\,Y_4$ be a similar arc on $P_2\,P_3$, and let $Y_5\,Y_6$ be a similar arc on $P_3\,P_1$. Some two of the points Y_1, Y_3, Y_5 must lie on the same arc $A'\,B'$ of W; suppose Y_1 and Y_3 do (Fig. 7). Now consider the simple closed curve

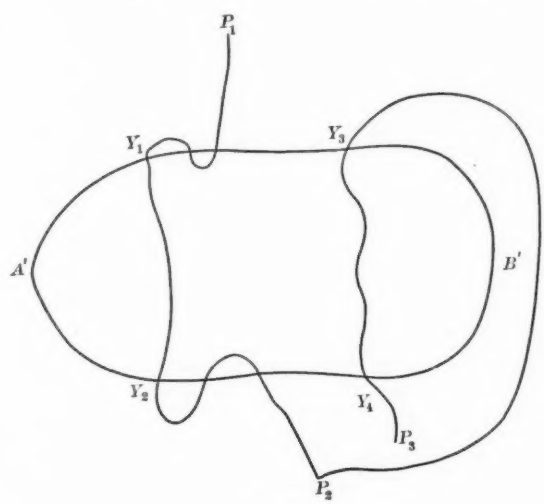

Fig. 7.

$\overline{Y} = Y_1\,Y_3\,Y_4\,Y_2\,Y_1$. Since it contains neither A' nor B' it encloses no point of either K_1 or K_2. Now there is no connected subset of M_2 lying on or within $\overline{Y}$ and having a point $\overline{X}_3$ on $Y_1\,Y_3$ and a point $\overline{X}_4$ on $Y_2\,Y_4$. For such a set would separate A' and B' on W and would divide the interior of W into two sets of which one contains K_1 and the other contains K_2. But K_1 and K_2 are connected within W by a set of points belonging to M_1. This set would have to cross the subset of M_2 lying within Y. But this is obviously impossible. Then by Theorem 2 there is a division of the points of M_2 that lie on or within $\overline{Y}$ into two mutually exclusive closed sets, Z_1 and Z_2, such that Z_1

contains all the points of M_2 that lie on $Y_1 Y_3$. Now all points of Z_1 can be enclosed in a finite number of circles not containing any point of $Y_3 Y_4 + Y_4 Y_2 + Y_2 Y_1$ or any point of Z_2 or M_1, and evidently there is an arc lying on or within $\overline{Y}$, a subset of these circles together with intervals of $Y_1 Y_3$, which contains no point of either M_1 or M_2. Then P_1 and P_2 are joined by an arc not containing any point of either M_1 or M_2. Since this is contrary to our supposition the theorem is proved.

University of Texas,
Austin, Tex.